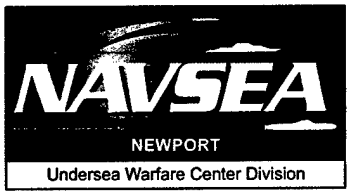
NUWC-NPT Technical Report 11,512
23 February 2004

Optimum Detection of Random Signal in Non-Gaussian Noise for Low Input Signal-to-Noise Ratio

Albert H. Nuttall
Surface Undersea Warfare Department



Naval Undersea Warfare Center Division
Newport, Rhode Island

PREFACE

The work described in this report was jointly funded under the Naval Undersea Warfare Center's "Large-N Strategic Initiative" project, principal investigator Clifford M. Curtis (Code 21), and the Office of Naval Research's "Large-N Discovery and Invention (D&I)" project, principal investigator Stephen G. Greineder (Code 2121).

The technical reviewer for this report was Stephen G. Greineder (Code 2121).

Reviewed and Approved: 23 February 2004

Donald A. Aker
Head (Acting), Surface Undersea Warfare Department

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	23 February 2004	

4. TITLE AND SUBTITLE

Optimum Detection of Random Signal in Non-Gaussian Noise for Low Input Signal-to-Noise Ratio

5. FUNDING NUMBERS

6. AUTHOR(S)

Albert H. Nuttall

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Undersea Warfare Center Division
1176 Howell Street
Newport, RI 02841-1708

8. PERFORMING ORGANIZATION REPORT NUMBER

TR 11,512

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
Ballston Centre Tower One
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for pubic release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Optimum detection of a weak stationary random signal in independent non-Gaussian noise requires knowledge of the first-order probability density function of the noise and the covariance function of the signal. More precisely, the first and second derivatives of the input noise probability density function must be known to realize the optimum processor. When these two derivatives must be estimated from a finite segment of noise-only data, a severe demand is placed on the amount of required data. Estimation of higher derivatives of histograms is not accomplished reliably without considerable amounts of data. The presence of heavy-tailed noise data exacerbates this issue. The situation improves when Gaussian noise is considered, mainly because the second derivative of the noise density is not relevant or required for Gaussian noise; all other noise densities must have this information to achieve optimum detection.

The samples of the random input signal process need not be taken at an independent rate. However, the covariance of the signal process must be known for optimum processing. The joint probability density function of the input signal is not required for low input signal-to-noise ratios. A simple example of a multipath signal is presented in this report that indicates the need for knowledge of the relative path strengths and the multipath delay time. Lack of knowledge of these parameters in this model requires multiple guesses at their values and parallel processors; the amount of degradation depends on the uncertainty of the medium characteristics.

14. SUBJECT TERMS

Optimum Detection
Probability Density Function
Random Signals
Heavy-Tailed Noise
Non-Gaussian Noise
Multipath Signals

15. NUMBER OF PAGES
48

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont'd)

LIST OF ABBREVIATIONS, ACRONYMS, AND SYMBOLS

$\mathbf{a}$	$K \times 1$ auxiliary vector, equation (13)
$a(x), b(x)$	Auxiliary functions, equation (7)
$c_{\mathbf{x}}(x)$	Cumulative distribution function of x, equation (38)
$c_{\mathbf{y}}(y)$	Cumulative distribution function of y, equation (38)
$\mathbf{d}_k$	Direct path component sample, equation (32)
EDF	Exceedance distribution function
$f, \tilde{f}$	Nonlinear transformation and its inverse, equation (38)
FO	First order, equation (4)
H_0, H_1	Signal absent, present hypotheses, equation (3)
k	Time index, equation (3)
K	Number of time samples, equation (3)
$L(\ )$	Likelihood ratio, equations (12), (13), (16)
LR	Likelihood ratio
m	Multipath delay, equation (32)
$\mathbf{n}_k$	Noise sample, equation (3)
$p_n(x)$	Probability density function of noise, equation (4)
$p_{nn}(\ )$	Joint noise density, equation (4)
$p_s(\ )$	Probability density function of signal, equation (5)
$p_{sn}(\ )$	Joint density of signal and noise, equation (4)
$p_{\mathbf{x}}(x)$	First-order density of $\boldsymbol{x}$, equation (38)

LIST OF ABBREVIATIONS, ACRONYMS, AND SYMBOLS (Cont'd)

$p_y(y)$	First-order density of y, equation (39)
PDF	Probability density function
Pr( )	Probability of ( ), equation (38)
$q(x)$	$\log p_n(x)$, equation (14)
$q_1 = q'(x)$	First derivative of q, equations (15) and (47)
$q_2 = q''(x)$	Second derivative of q, equations (15) and (47)
r	Signal-to-noise power ratio, equation (25)
R_{sn}	Signal-plus-noise covariance matrix, equation (22)
ROC	Receiver operating characteristic
RV	Random variable
$\mathbf{s}_k$	Signal sample, equation (3)
SNR	Signal-to-noise ratio
$t(y)$	Auxiliary variable, equations (44) and (46)
$\mathbf{x}$	Gaussian random variable, equation (37)
$\mathbf{x}_k$	Sample of received data, equation (3)
$\mathbf{X}$	$K \times 1$ data vector, equation (19)
$\mathbf{y}$	Transformed random variable, equations (37) and (38)
α	Multipath component relative strength, equation (32)
β	Noise parameter, equations (28), (29), (31)
ρ	Normalized signal covariance matrix, equation (13)
$\widetilde{\rho}$	Modified signal covariance matrix, equation (13)
ρ_{kj}	Normalized signal covariance element, equation (10)
$\widetilde{\rho}_{kj}$	Modified signal covariance element, equation (12)
σ_d	Direct path signal standard deviation, equations (33), (35)
σ_s	Signal standard deviation, equation (10)
σ_n	Noise standard deviation, equation (18)

boldface	Random variable
prime	Derivative
superscript T	Transpose, equation (13)
superscript -1	Inverse matrix, equation (23)

OPTIMUM DETECTION OF RANDOM SIGNAL IN NON-GAUSSIAN NOISE FOR LOW INPUT SIGNAL-TO-NOISE RATIO

INTRODUCTION

A finite-time observation is made of a weak signal (if present) in a non-Gaussian noise background. Optimum detection consists of computing the likelihood ratio (LR) of the observed waveform samples and comparing the LR with a fixed threshold, which enables a statement about signal presence versus absence in that time interval. The random noise samples are denoted by $\{\mathbf{n}_k\}$ for $k = 1{:}K$. It is presumed that these noise samples are statistically independent with common first-order probability density function (PDF) $p_n(x)$. If the signal is present, its samples are $\{s_k\}$.

For a deterministic signal $\{s_k\}$, the LR and log LR take the forms

$$\mathbf{LR} = \prod_{k=1}^{K} \frac{p_n(\mathbf{x}_k - s_k)}{p_n(\mathbf{x}_k)}, \quad \log \mathbf{LR} = \sum_{k=1}^{K} [q(\mathbf{x}_k - s_k) - q(\mathbf{x}_k)], \tag{1}$$

where $\{\mathbf{x}_k\}$ is the received data, and $q(x) = \log p_n(x)$. For low input signal-to-noise ratio (SNR), the approximation

$$\log \mathbf{LR} \cong -\sum_{k=1}^{K} q'(\mathbf{x}_k)\, s_k = -\sum_{k=1}^{K} \frac{p_n'(\mathbf{x}_k)}{p_n(\mathbf{x}_k)}\, s_k \tag{2}$$

follows. Thus, for low SNR, the dominant term in the optimum processor depends on the derivative of the q function and is linear with the signal samples $\{s_k\}$.

On the other hand, for a stationary random signal with *zero mean*, the optimum processor for low input SNR must depend on higher-order statistics of the input signal. At the very least, the covariance of the input signal will be required or must be estimated. Also, it can be expected that higher-order information about the non-Gaussian noise PDF $p_n(x)$ will be required in addition to $q'(x)$. This is indeed the case. To make reliable decisions about signal presence or absence, these more stringent information requirements for the random signal case require additional data on both the received signal and the noise.

DERIVATION OF LIKELIHOOD RATIO PROCESSOR

Samples $\{\mathbf{x}_k\}$ for $k = 1{:}K$ are available. Under the hypotheses

$$\begin{aligned} H_0 &: \mathbf{x}_k = \mathbf{n}_k, \\ H_1 &: \mathbf{x}_k = \mathbf{s}_k + \mathbf{n}_k, \end{aligned} \tag{3}$$

the optimum detector of signal presence versus absence is desired. Noise sequence $\{\mathbf{n}_k\}$ consists of independent samples, with known first-order (FO) PDF $p_n(x)$, which is arbitrary. Random signal sequence $\{\mathbf{s}_k\}$ is characterized by joint K-dimensional PDF $p_s(u_1,\ldots,u_K)$. The optimum LR detector forms the ratio

$$\mathbf{LR} = \frac{p_{sn}(\mathbf{x}_1,\ldots,\mathbf{x}_K)}{p_{nn}(\mathbf{x}_1,\ldots,\mathbf{x}_K)} = \frac{p_{sn}(\mathbf{x}_1,\ldots,\mathbf{x}_K)}{\prod_{k=1}^{K} p_n(\mathbf{x}_k)} \tag{4}$$

for comparison with a fixed threshold.

The numerator of equation (4) is given by K-dimensional convolution:

$$p_{sn}(\mathbf{x}_1,\ldots,\mathbf{x}_K) = \int\cdots\int du_1\cdots du_K \; p_s(u_1,\ldots,u_K)\, p_n(\mathbf{x}_1 - u_1)\cdots p_n(\mathbf{x}_K - u_K). \tag{5}$$

For low input SNR and zero-mean processes, the joint PDF $p_s(u_1,\ldots,u_K)$ is concentrated near the origin of K-dimensional u-space, relative to the spread of the remaining FO noise PDFs $\{p_n(\mathbf{x}_k - u_k)\}$. Therefore, a typical noise-PDF term in equation (5) can be expanded near $u_k = 0$ according to

$$p_n(\mathbf{x}_k - u_k) \cong p_n(\mathbf{x}_k) - u_k\, p_n'(\mathbf{x}_k) + \tfrac{1}{2}u_k^2\, p_n''(\mathbf{x}_k) \quad \text{for} \quad k = 1{:}K. \tag{6}$$

This expansion is to second order in u_k.

Two auxiliary functions are defined as

$$a(x) = -\frac{p_n'(x)}{p_n(x)}, \quad b(x) = \frac{p_n''(x)}{p_n(x)}. \tag{7}$$

Then, equation (6) becomes

$$p_n(\mathbf{x}_k - u_k) \cong p_n(\mathbf{x}_k)\,[1 + u_k\; a(\mathbf{x}_k) + \tfrac{1}{2} u_k^2\; b(\mathbf{x}_k)] \quad \text{for} \quad k = 1:K. \tag{8}$$

Substitution of this expansion in equation (5) and retention of terms through second order in $\{u_k\}$, which is consistent with second-order expansion (6), yields

$$p_{sn}(\mathbf{x}_1,\ldots,\mathbf{x}_K) \cong \prod_k \{p_n(\mathbf{x}_k)\} \int\cdots\int du_1 \cdots du_K\; p_s(u_1,\ldots,u_K)$$

$$\times \left[1 + \sum_k u_k\; a(\mathbf{x}_k) + \tfrac{1}{2}\sum_k u_k^2\; b(\mathbf{x}_k) + \tfrac{1}{2}\sum_{k\neq j} u_k\; u_j\; a(\mathbf{x}_k)\; a(\mathbf{x}_j)\right]. \tag{9}$$

The three types of integrals in equation (9) can be evaluated as

$$\begin{aligned}
&\int\cdots\int du_1 \cdots du_K\; p_s(u_1,\ldots,u_K)\; u_k = \int du_k\; p_{s1}(u_k)\; u_k = 0, \\
&\int\cdots\int du_1 \cdots du_K\; p_s(u_1,\ldots,u_K)\; u_k^2 = \int du_k\; p_{s1}(u_k)\; u_k^2 = \sigma_s^2, \\
&\int\cdots\int du_1 \cdots du_K\; p_s(u_1,\ldots,u_K)\; u_k\; u_j = \iint du_k\; du_j\; p_{s2}(u_k,u_j)\; u_k\; u_j = \sigma_s^2\; \rho_{kj},
\end{aligned} \tag{10}$$

where σ_s^2 is the common signal variance of samples $\{\mathbf{s}_k\}$, and ρ_{kj} is the normalized signal covariance coefficient of $\mathbf{s}_k$ and $\mathbf{s}_j$.

The use of equation (10) in equation (9), followed by substitution in equation (4), leads to the approximation

$$\mathbf{LR} \cong 1 + \tfrac{1}{2}\sigma_s^2 \sum_k b(\mathbf{x}_k) + \tfrac{1}{2}\sigma_s^2 \sum_{k\neq j} \rho_{kj}\; a(\mathbf{x}_k)\; a(\mathbf{x}_j) \text{ for low input SNR.} \tag{11}$$

Since data-independent additive constants and positive scalars do not affect the receiver operating characteristics (ROCs) of a processor, an equivalent processor to equation (11) is

$$L(\mathbf{x}_1,\ldots,\mathbf{x}_K) \equiv \sum_{k=1}^{K} b(\mathbf{x}_k) + \sum_{k,j=1}^{K} \tilde{\rho}_{kj}\; a(\mathbf{x}_k)\; a(\mathbf{x}_j), \tag{12}$$

where modified covariance matrix $\tilde{\rho} = [\tilde{\rho}_{kj}]$ is obtained from $K \times K$ matrix $\rho = [\rho_{kj}]$ by replacing all the unity diagonal elements with zeros. This modification enables equation (12) to be expressed in the compact form

$$L(\mathbf{x}_1,\ldots,\mathbf{x}_K) = \sum_{k=1}^{K} b(\mathbf{x}_k) + \mathbf{a}^T\; \tilde{\rho}\; \mathbf{a}, \tag{13}$$

where $K \times 1$ vector $\mathbf{a} = [a(\mathbf{x}_1) \cdots a(\mathbf{x}_K)]^T$. This is the (approximate) optimum detector for low input SNR and *any* joint signal statistics. Although the signal level is not involved in equation (13), the normalized signal covariance matrix ρ must be known to employ equation (13).

An alternative form for the optimum processor is obtained by defining

$$q(x) = \log p_n(x). \tag{14}$$

Then, using equation (7),

$$q'(x) = \frac{p_n'(x)}{p_n(x)} = -a(x), \quad q''(x) = \frac{p_n(x)\, p_n''(x) - p_n'(x)^2}{p_n(x)^2} = b(x) - a(x)^2. \tag{15}$$

The optimum processor in equation (12) can now be expressed as

$$\begin{aligned} L(\mathbf{x}_1, \ldots, \mathbf{x}_K) &= \sum_{k=1}^{K} b(\mathbf{x}_k) - \sum_{k=1}^{K} a(\mathbf{x}_k)^2 + \sum_{k,j=1}^{K} \rho_{kj}\, a(\mathbf{x}_k)\, a(\mathbf{x}_j) \\ &= \sum_{k=1}^{K} q''(\mathbf{x}_k) + \sum_{k,j=1}^{K} \rho_{kj}\, q'(\mathbf{x}_k)\, q'(\mathbf{x}_j). \end{aligned} \tag{16}$$

An arbitrary constant can be added to $q''(x)$ without affecting the performance of this processor.

As a special case, if the signal samples $\{\mathbf{s}_k\}$ are uncorrelated, then matrix $\tilde{\rho} = 0$, and equation (13) reduces to

$$L_0(\mathbf{x}_1, \ldots, \mathbf{x}_K) = \sum_{k=1}^{K} b(\mathbf{x}_k) = \sum_{k=1}^{K} \frac{p_n''(\mathbf{x}_k)}{p_n(\mathbf{x}_k)}. \tag{17}$$

Nonlinearities $q'(x)$ and $a(x)$ are not involved at all in this processor. However, more generally, when the signal samples are correlated, then matrix $\tilde{\rho} \neq 0$ and the last term in equations (12) and (13) allows for interaction *between* the $\{\mathbf{x}_k\}$ terms by means of the nonlinear transformation $a(x) = -\, p_n'(x)/p_n(x)$.

The noise samples $\{\mathbf{n}_k\}$ have been taken statistically independent in the derivation above. If the sampling rate used to generate the data is increased, the noise samples will become statistically dependent on each other. This effect is investigated in appendix A for a simple example. It is shown that the deflection criterion is insignificantly improved through an increased sampling rate. Sampling at the highest rate that still results in uncorrelated samples will extract all the detection capability that a given data segment length contains. Higher sampling rates just require more data processing, with little if any improvement in performance.

GAUSSIAN NOISE PDF

$$
\begin{aligned}
&p_n(x) = \frac{1}{\sqrt{2\pi}\sigma_n}\exp\left(-\frac{x^2}{2\sigma_n^2}\right),\\
&p_n'(x) = \frac{-x}{\sigma_n^2}p_n(x),\ \ p_n''(x) = \left(\frac{x^2}{\sigma_n^4} - \frac{1}{\sigma_n^2}\right)p_n(x),\\
&a(x) = \frac{x}{\sigma_n^2},\ \ b(x) = \frac{1}{\sigma_n^4}\left(x^2 - \sigma_n^2\right),\\
&L(\mathbf{x}_1,\ldots,\mathbf{x}_K) = \frac{1}{\sigma_n^4}\left[\sum_k\left(\mathbf{x}_k^2 - \sigma_n^2\right) + \sum_{k\neq j}\rho_{kj}\,\mathbf{x}_k\,\mathbf{x}_j\right].
\end{aligned}
\tag{18}
$$

Discarding additive constants and positive scalars, an equivalent processor is

$$
\sum_k \mathbf{x}_k^2 + \sum_{k\neq j}\rho_{kj}\,\mathbf{x}_k\,\mathbf{x}_j = \sum_{k,j=1}^{K}\rho_{kj}\,\mathbf{x}_k\,\mathbf{x}_j = \mathbf{X}^T\rho\,\mathbf{X},
\tag{19}
$$

where $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_K]^T$. This is a "generalized" energy detector, weighted by the normalized signal covariance matrix ρ.

Alternatively, from equations (14) through (16),

$$
\begin{aligned}
&q(x) = -\frac{x^2}{2\sigma_n^2} - \log\left(\sqrt{2\pi}\sigma_n\right),\ \ q'(x) = -\frac{x}{\sigma_n^2},\ \ q''(x) = -\frac{1}{\sigma_n^2},\\
&L(\mathbf{x}_1,\ldots,\mathbf{x}_K) = -\frac{K}{\sigma_n^2} + \frac{1}{\sigma_n^4}\sum_{k,j=1}^{K}\rho_{kj}\,\mathbf{x}_k\,\mathbf{x}_j,
\end{aligned}
\tag{20}
$$

which is tantamount to processor (19). The $q''(x)$ term has *no effect* on the performance of the processor for the case of Gaussian noise.

The results in equations (19) and (20) hold for low input SNR, σ_s^2/σ_n^2, and *any* signal joint PDF $p_s(u_1,\ldots,u_K)$. It is informative to compare these results with the exact LR processor for Gaussian signal and Gaussian noise. The latter processor is, from equation (4) and discarding a positive scalar,

$$
\mathbf{LR} = \exp\left[-\frac{1}{2}\mathbf{X}^T R_{sn}^{-1}\mathbf{X} + \frac{1}{2\sigma_n^2}\mathbf{X}^T\mathbf{X}\right] = \exp\left[\frac{1}{2}\mathbf{X}^T\left(\frac{1}{\sigma_n^2}I - R_{sn}^{-1}\right)\mathbf{X}\right],
\tag{21}
$$

where signal-plus-noise covariance matrix

$$R_{sn} = \sigma_n^2 I + \sigma_s^2 \rho = \sigma_n^2 \left(I + \frac{\sigma_s^2}{\sigma_n^2} \rho \right). \tag{22}$$

For low input power SNR, $\sigma_s^2/\sigma_n^2 << 1$, the inverse of matrix R_{sn} is

$$R_{sn}^{-1} \cong \frac{1}{\sigma_n^2} \left(I - \frac{\sigma_s^2}{\sigma_n^2} \rho \right). \tag{23}$$

Substituting in equation (21) yields

$$\mathbf{LR} \cong \exp\left[\frac{1}{2} \frac{\sigma_s^2}{\sigma_n^4} \mathbf{X}^T \rho \, \mathbf{X} \right]. \tag{24}$$

This is simply a monotonic transformation of equivalent processors (19) and (20); this equivalence holds under approximation (23).

More precisely, using the identity

$$I - (I + r\rho)^{-1} = r\rho \, (I + r\rho)^{-1}, \; r = \sigma_s^2 / \sigma_n^2 < 1, \tag{25}$$

equation (21) becomes exactly

$$\mathbf{LR} = \exp\left[\frac{1}{2} \frac{\sigma_s^2}{\sigma_n^4} \mathbf{X}^T \rho \, (I + r\rho)^{-1} \mathbf{X} \right]. \tag{26}$$

An equivalent exact LR processor is, for this Gaussian noise PDF,

$$\begin{aligned} L(\mathbf{x}_1, \ldots, \mathbf{x}_K) &= \mathbf{X}^T \rho \, (I + r\rho)^{-1} \mathbf{X} = \mathbf{X}^T \rho \, (I - r\rho + r^2\rho^2 - \cdots) \mathbf{X} \\ &= \mathbf{X}^T \rho \, \mathbf{X} - r \, \mathbf{X}^T \rho^2 \, \mathbf{X} + r^2 \, \mathbf{X}^T \rho^3 \, \mathbf{X} - \cdots . \end{aligned} \tag{27}$$

Terms beyond the first involve additional factors containing powers of input SNR $r = \sigma_s^2/\sigma_n^2$ and matrix ρ, both of which allow for dropping these terms for low input SNR with negligible effect on detection performance.

Expansion (9) was conducted only through second order. In appendix B, the third- and fourth-order expansions are derived for those interested in pursuing these higher-order terms numerically.

EXPONENTIAL NOISE PDF

$$
\begin{aligned}
&p_n(x) = \frac{1}{2\beta}\exp\left(-\frac{|x|}{\beta}\right),\ \ \sigma_n^2 = 2\beta^2,\\
&q(x) = -\frac{|x|}{\beta} - \log(2\beta),\\
&q'(x) = \left\{\begin{matrix} -1/\beta & \text{for } x > 0 \\ 1/\beta & \text{for } x < 0 \end{matrix}\right\} = -\frac{1}{\beta}\operatorname{sgn}(x),\ \ \beta = \sigma_n/\sqrt{2},\\
&q''(x) = -\frac{2}{\beta}\,\delta(x).
\end{aligned}
\tag{28}
$$

The difficulty in this example is due to noise PDF $p_n(x)$ not possessing a derivative at $x = 0$. The last term in equation (16) becomes $\frac{1}{\beta^2}\sum_{kj}\rho_{kj}\ \operatorname{sgn}(\mathbf{x}_k)\ \operatorname{sgn}(\mathbf{x}_j)$, which does not appear to have an immediate problem. However, the first term involving $q''(x)$ obviously needs special treatment and is taken up in appendix C.

CAUCHY NOISE PDF

$$
\begin{aligned}
&p_n(x) = \frac{\beta}{\pi}\frac{1}{x^2+\beta^2},\ \ p_n'(x) = \frac{\beta}{\pi}\frac{-2x}{\left(x^2+\beta^2\right)^2},\ \ p_n''(x) = \frac{\beta}{\pi}\frac{6x^2-2\beta^2}{\left(x^2+\beta^2\right)^3},\\
&a(x) = \frac{2x}{x^2+\beta^2},\ \ b(x) = \frac{6x^2-2\beta^2}{\left(x^2+\beta^2\right)^2},\ \ q'(x) = -a(x),\ \ q''(x) = \frac{2(x^2-\beta^2)}{\left(x^2+\beta^2\right)^2}.
\end{aligned}
\tag{29}
$$

Large $\mathbf{x}_k$ values are suppressed by both components, q' and q'', of the optimum processor:

$$
q'(\mathbf{x}_k)\,q'(\mathbf{x}_j) \sim \frac{4}{\mathbf{x}_k\,\mathbf{x}_j},\ \ q''(\mathbf{x}_k) \sim \frac{2}{\mathbf{x}_k^2}.
\tag{30}
$$

ALTERNATIVE NOISE PDF $(\alpha = \beta^3\sqrt{2}/\pi)$

$$
\begin{aligned}
&p_n(x) = \frac{\alpha}{x^4+\beta^4},\ \ p_n'(x) = \alpha\frac{-4x^3}{\left(x^4+\beta^4\right)^2},\ \ p_n''(x) = \alpha\,4x^2\frac{5x^4-3\beta^4}{\left(x^4+\beta^4\right)^3},\\
&a(x) = \frac{4x^3}{x^4+\beta^4},\ \ b(x) = 4x^2\frac{5x^4-3\beta^4}{\left(x^4+\beta^4\right)^2},\ \ q''(x) = 4x^2\frac{x^4-3\beta^4}{\left(x^4+\beta^4\right)^2}.
\end{aligned}
\tag{31}
$$

Again, large $\mathbf{x}_k$ values are suppressed in a similar fashion to equation (30). But, small $\mathbf{x}_k$ values are also suppressed by $a(x)$, which has cubic behavior near the origin. Odd function $a(x)$ has a peak at $x = 3^{1/4}\beta$ of value $3^{3/4}/\beta$.

Some additional impulsive noise models are given in appendix D.

MULTIPATH EXAMPLE: GENERAL NOISE STATISTICS

Letting the received signal contain a direct path component and a delayed (surface bounce) multipath component:

$$\mathbf{s}_k = \mathbf{d}_k + \alpha\, \mathbf{d}_{k-m}, \quad m > 0, \tag{32}$$

where zero-mean sequence $\{\mathbf{d}_k\}$ consists of uncorrelated samples,

$$E\{\mathbf{d}_k\, \mathbf{d}_j\} = \sigma_d^2\, \delta(k, j). \tag{33}$$

Then, signal covariance

$$E\{\mathbf{s}_k\, \mathbf{s}_j\} = \sigma_d^2\, [\delta(k, j) + \alpha\, \delta(k-m, j) + \alpha\, \delta(j-m, k) + \alpha^2\, \delta(k-m, j-m)], \tag{34}$$

leading to total received signal power $\sigma_s^2 = \sigma_d^2\, (1+\alpha^2)$ and

$$\rho = \begin{bmatrix} 1 & 0 & \cdots & 0 & \beta & 0 & \cdots \\ 0 & 1 & 0 & & 0 & \beta & 0 \\ \vdots & 0 & \ddots & \ddots & & \ddots & \ddots \\ 0 & & \ddots & & & & \\ \beta & 0 & & & & & \\ 0 & \beta & \ddots & & & & \\ \vdots & 0 & \ddots & & & & \end{bmatrix}, \quad \beta = \frac{\alpha}{1+\alpha^2}, \tag{35}$$

where value $\beta = \alpha/(1+\alpha^2)$ appears in the super- and sub-diagonals that are m removed from the main diagonal. Optimum processor (12) or (13) now takes the form

$$L(\mathbf{x}_1, \ldots, \mathbf{x}_K) = \sum_{k=1}^{K} b(\mathbf{x}_k) + \frac{2\alpha}{1+\alpha^2} \sum_{k=1}^{K-m} a(\mathbf{x}_k)\, a(\mathbf{x}_{k+m}) \tag{36}$$

for low input SNR, where nonlinear transformations $a(x)$ and $b(x)$ are given by equation (7). The first term in equation (36) is a sum of samples $\{\mathbf{x}_k\}$ subject to nonlinearity $b(x)$, while the second term is an autocorrelation (at delay m) of distorted samples $\{a(\mathbf{x}_k)\}$. Realization of processor (36) requires knowledge of scaling α and delay m in model (32). Lack of this

information will require several guesses of both parameters and repeated evaluation of the second term in equation (36) to realize a maximum of $L(\mathbf{x}_1,...,\mathbf{x}_K)$. However, the latter sum in equation (36) does not, by itself, require knowledge of α, while its scale factor depends solely on α. This observation can be used to simplify the computation of equation (36) for multiple guesses of α and m.

An alternative expression for the optimum processor in equation (36) is available upon use of equation (15); namely,

$$L(\mathbf{x}_1,\ldots,\mathbf{x}_K) = \sum_{k=1}^{K} q''(\mathbf{x}_k) + \sum_{k=1}^{K} q'(\mathbf{x}_k)^2 + \frac{2\alpha}{1+\alpha^2} \sum_{k=1}^{K-m} q'(\mathbf{x}_k)\, q'(\mathbf{x}_{k+m}).$$

If signal delay information m and signal strength information α are unknown, a suboptimum processor is available upon dropping the last term:

$$\tilde{L}(\mathbf{x}_1,\ldots,\mathbf{x}_K) \equiv \sum_{k=1}^{K} q''(\mathbf{x}_k) + \sum_{k=1}^{K} q'(\mathbf{x}_k)^2.$$

ESTIMATION OF OPTIMUM NONLINEARITIES FROM DATA

To implement the optimum processor in equations (13) or (16), it is necessary to determine the nonlinearities $q'(x)$ and $q''(x)$. When an analytic expression is available for noise PDF $p_n(x)$, as in equations (18), (28), (29), and (31), the appropriate derivatives can be taken and used in equation (15). This information is needed in equation (16) in addition to the signal covariance matrix ρ.

However, when an analytic noise PDF is not available, an estimate of the noise PDF must be made from a finite length of noise-alone measured data $\{\mathbf{n}_k\}$. Furthermore, two derivatives of the noise PDF must be extracted somehow from this finite data without exacerbating the effect of taking differences of noisy estimates. The method adopted here is to obtain a histogram from the measured data, fit it by means of a spline with several continuous derivatives, and then take two derivatives of the spline fit. In particular, the "spaps" procedure of MATLAB was used for all the following examples.

As a first example, 100,000 independent normalized Gaussian RVs were generated and subjected to the spaps procedure. The results for the three functions $q = q(x)$, $q_1 = q'(x)$, and $q_2 = q''(x)$ are shown in figure 1. An estimate for each function is displayed, as well as the exact nonlinearity from equation (20). The agreement among the three pairs is excellent over the full range of ±4.25 standard deviations plotted. The corresponding results for the smaller sample size of 10,000 data points is given in figure 2; the deviations between the exact and estimated results are now obvious but not excessive. Still smaller sample sizes lead to considerably worsened estimates of the three q functions.

The second example of interest is a non-Gaussian RV obtained by means of the nonlinear transformation

$$\mathbf{y} = \mathbf{x} + a\,\mathbf{x}^3, \tag{37}$$

where $\mathbf{x}$ is a Gaussian RV. Scale factor a is set equal to 0.1. For a unit-variance RV $\mathbf{x}$, the non-Gaussian RV $\mathbf{y}$ has the exact q functions displayed in figure 3; the derivations as well as a program for these functions are presented in appendix E. The heavier tails of this PDF, relative to the Gaussian case, are evident in the upper q function in figure 3, which was plotted over the range of ±5 standard deviations. The q_2 function has a pronounced dip of substantial value at the origin, while the q_1 function behaves like a limiter for input amplitudes greater than unity.

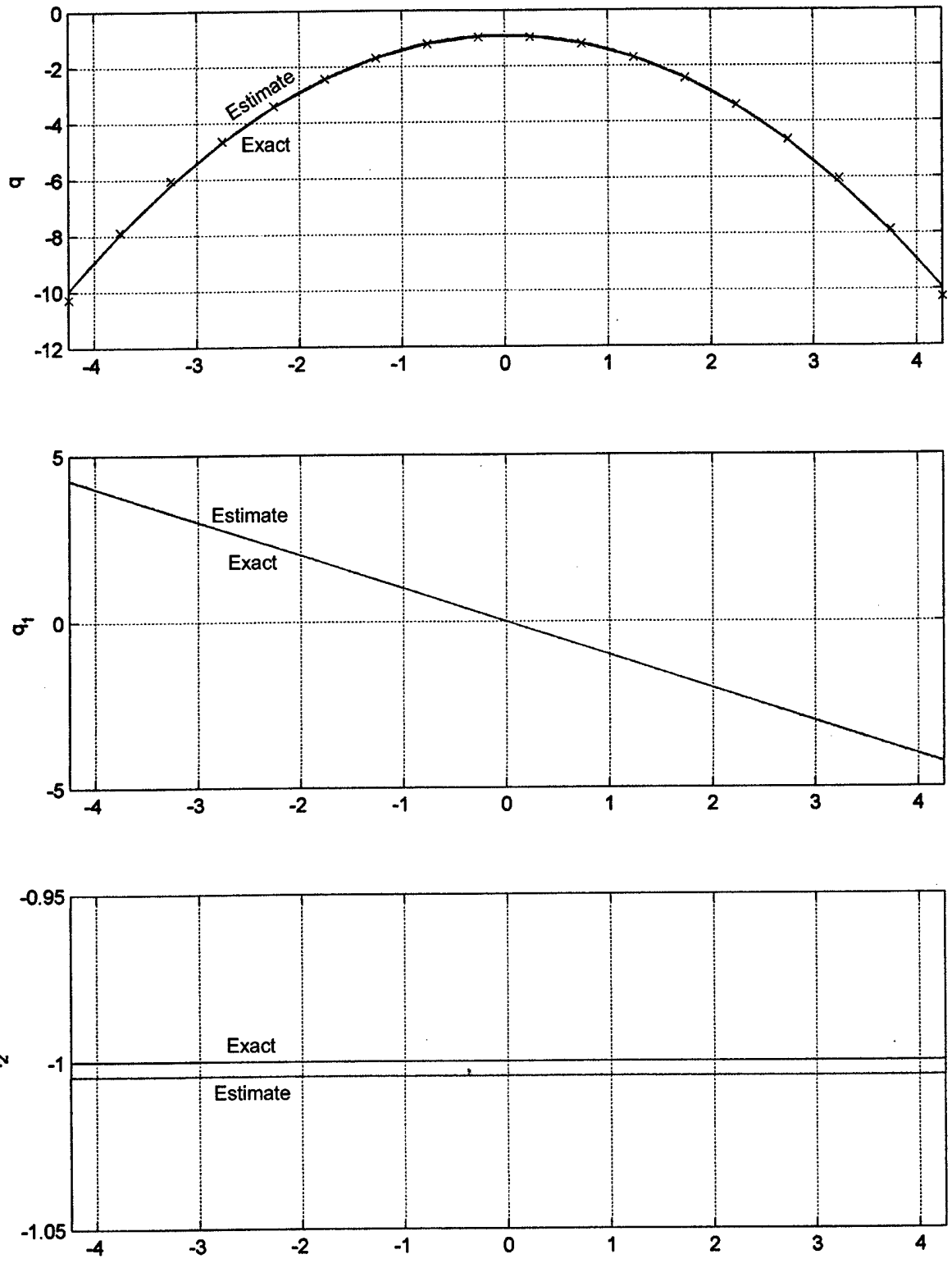


Figure 1. q Functions for 100,000 Gaussian Data Points

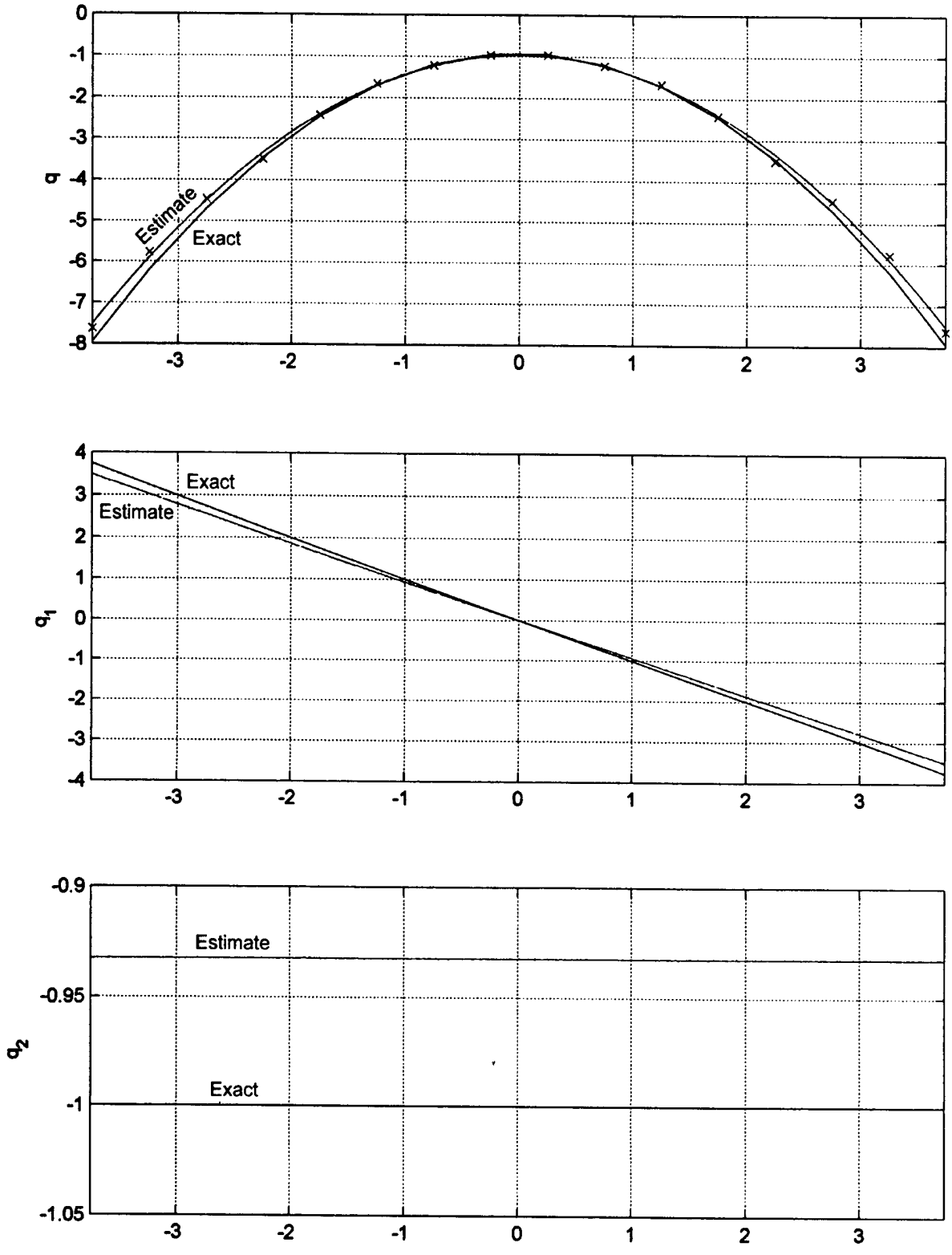


Figure 2. q Functions for 10,000 Gaussian Data Points

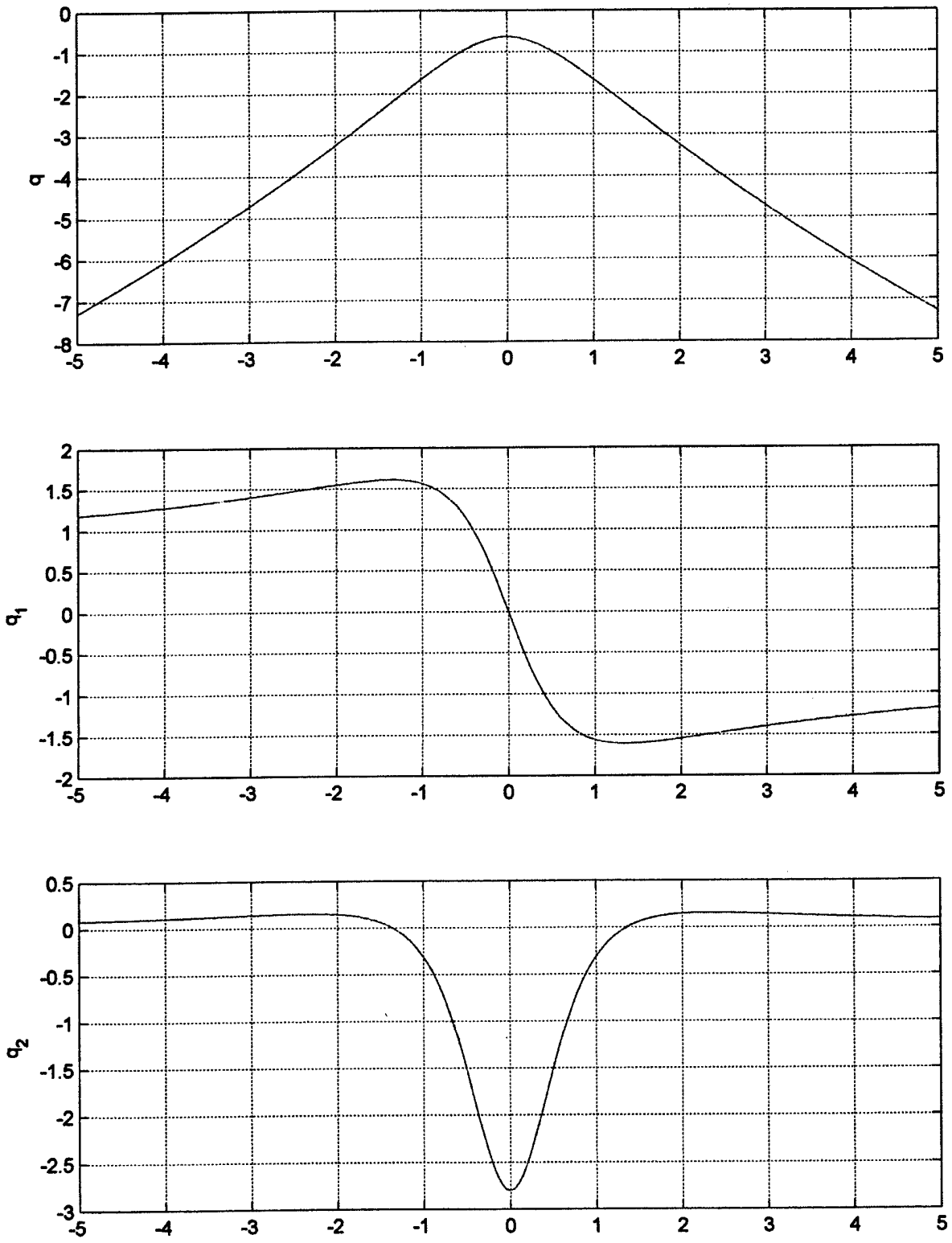


Figure 3. Exact q Functions for a = 0.1

Figures 4, 5, and 6 illustrate the type of stability and accuracy that can be attained from independent sample sizes of 1e7, 1e6, and 1e5, respectively, when the PDF and q functions must be estimated from finite data segments. The results in figure 4 for 1e7 samples agree very well with the exact results in figure 3, except on the edges near ±5 standard deviations. This edge effect is to be expected because of the rarity of encountering these values and the concurrent instability of the histogram estimate.

The results in figures 5 and 6 display much poorer agreement between estimation and truth, even though the two upper q functions appear to be fairly well fitted by the spline procedure. The problem lies in getting the higher-order derivatives, which accentuate the minor deviations in the fits. This non-Gaussian example illustrates that a considerable amount of data is going to be required if reliable estimates of the q_1 function, and especially of the q_2 function, are to be obtained directly from the data. This is not an unexpected result; estimation of higher-order derivatives from noisy data will always be difficult to accomplish successfully unless a very significant amount of independent data is available and is carefully processed.

Comparison of figures 1 and 6 also reveals that estimation of the q functions for the non-Gaussian noise case is considerably less stable; equivalent accuracy in the estimated results for the non-Gaussian noise case will require more data than for the Gaussian noise case. Furthermore, lack of sufficient data may preclude decent estimation of the q_2 function, with the attendant degradation of performance of the resultant suboptimum processor. Observe that for the Gaussian noise case, the q_2 function is a constant and can be totally *ignored* in the optimum processor. This is not the case for any non-Gaussian noise situation, where the q_2 contribution can be significant in making the proper detection decision.

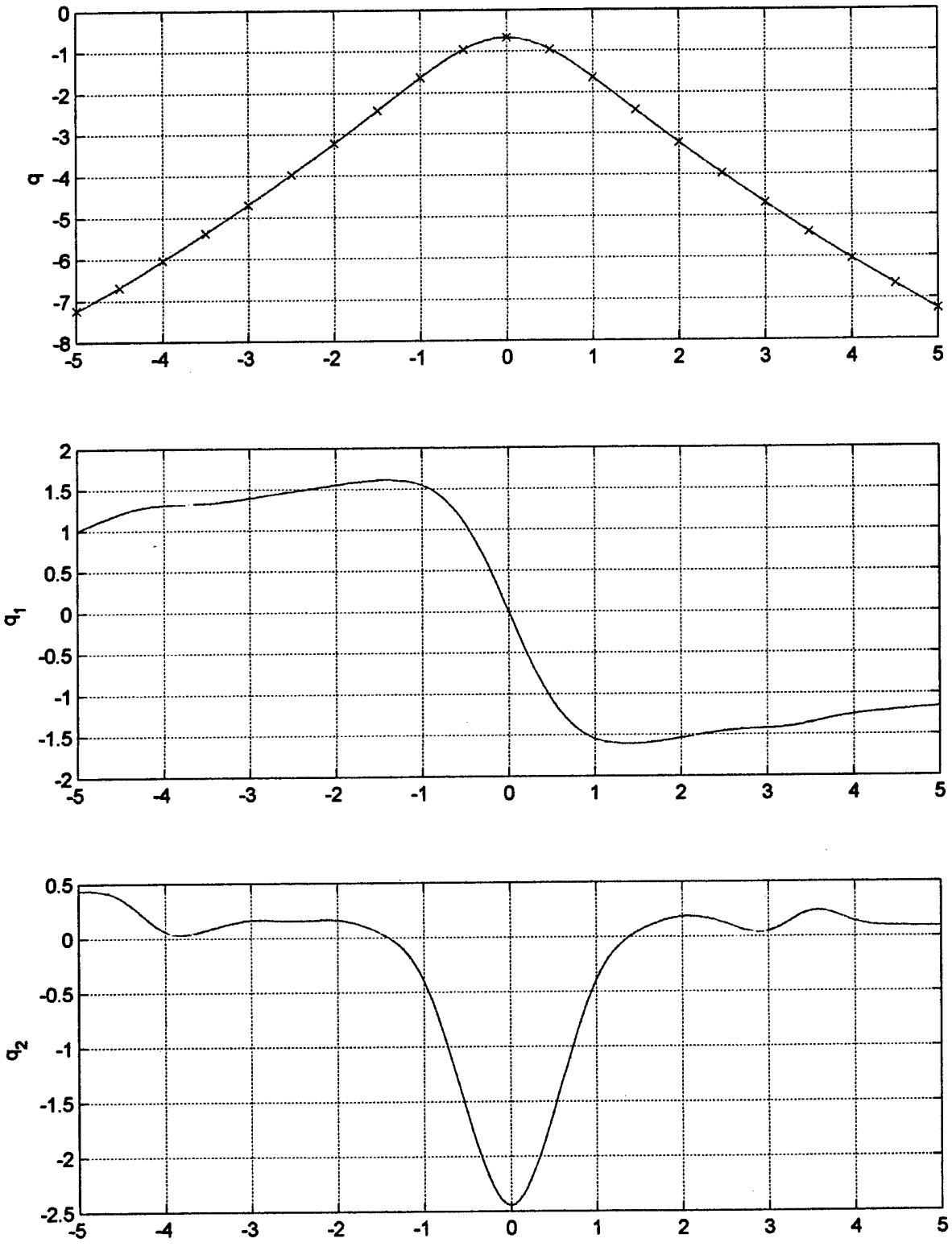


Figure 4. Estimated q Functions for a = 0.1, 1e7 Data Points

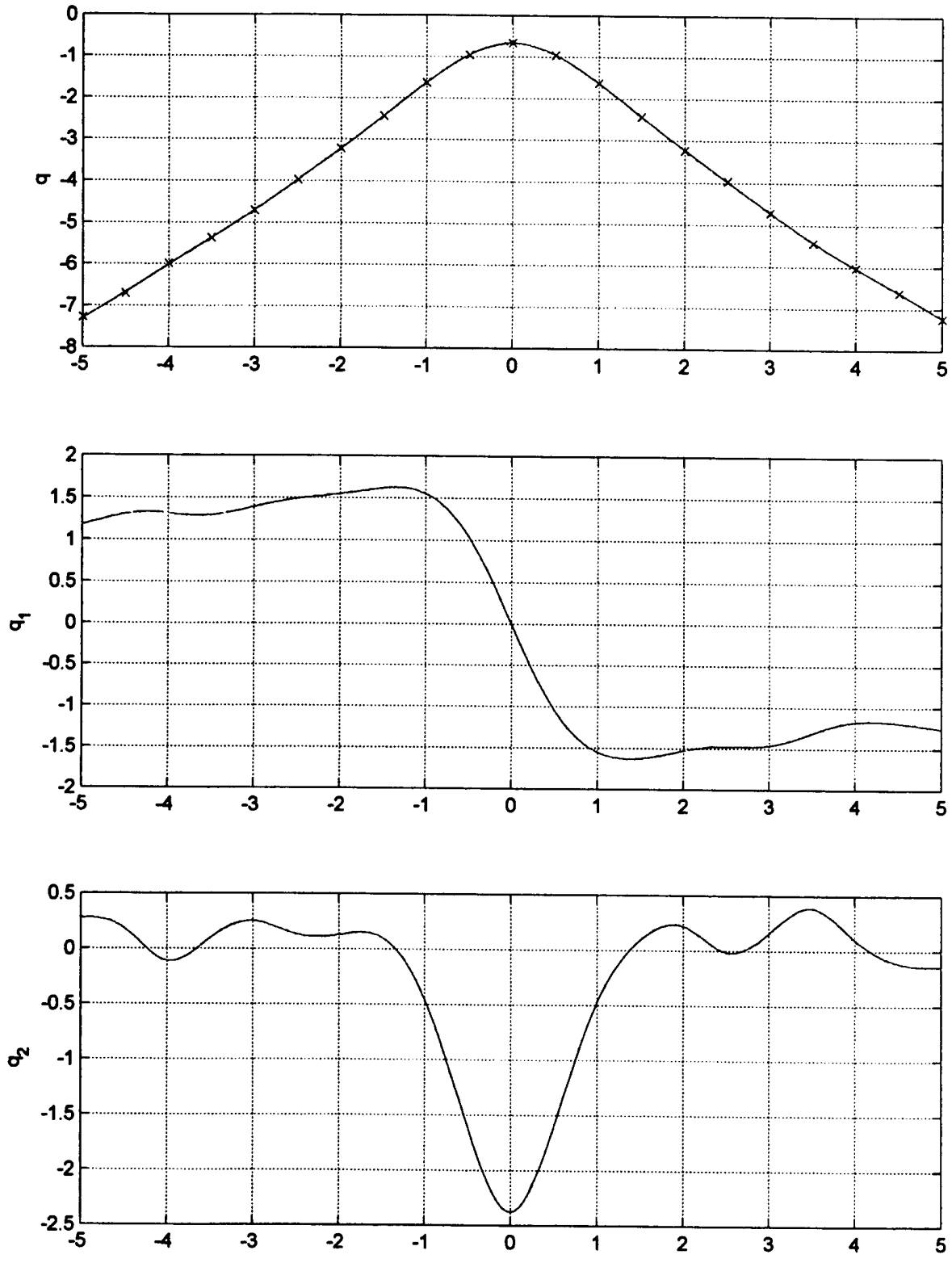


Figure 5. Estimated q Functions for a = 0.1, 1e6 Data Points

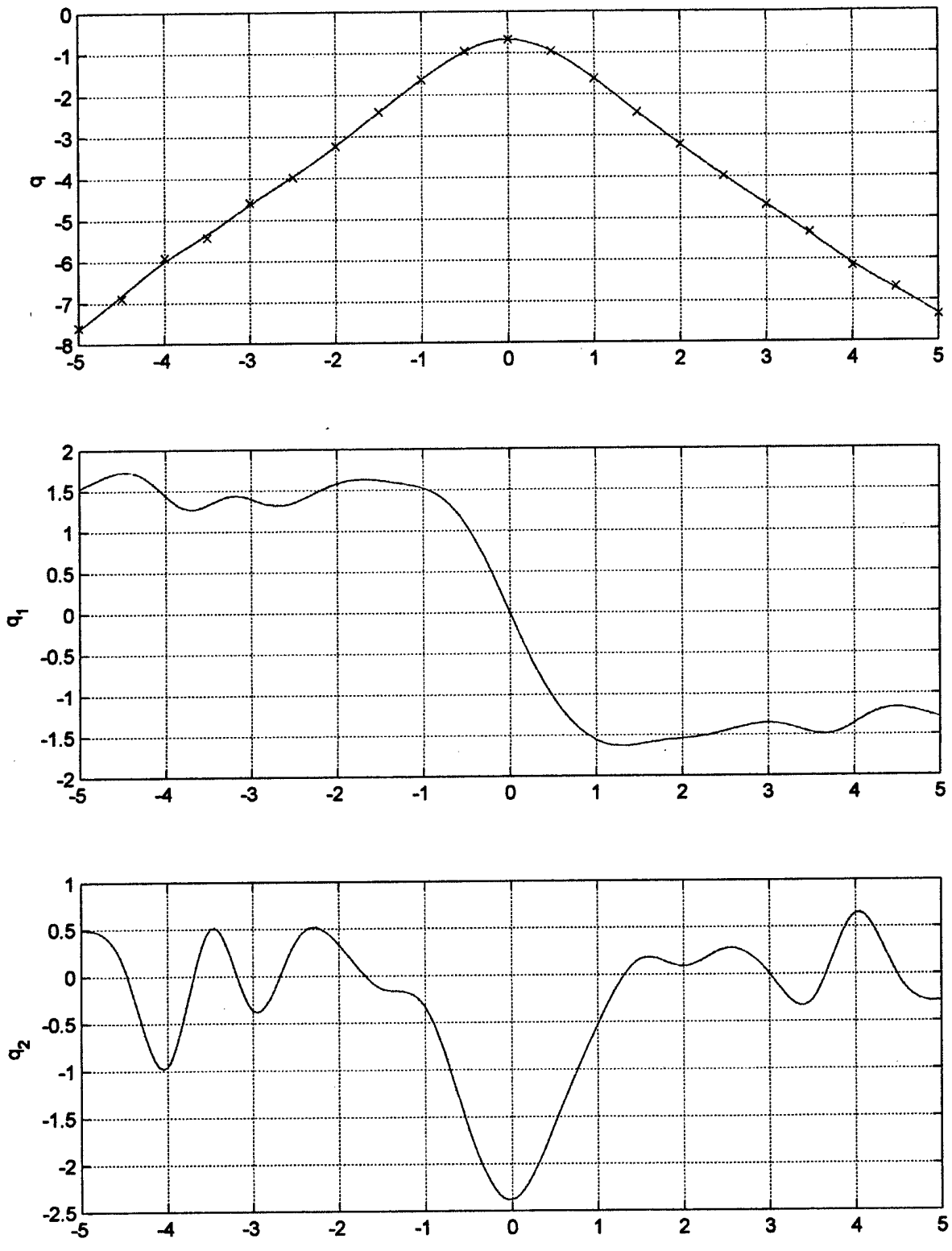


Figure 6. Estimated q Functions for a = 0.1, 1e5 Data Points

PERFORMANCE OF SEVERAL NONLINEAR PROCESSORS

Suppose that RV **x**, which has FO PDF $p_{\mathbf{x}}(x)$, is nonlinearly transformed by monotonically increasing function f, yielding RV $\mathbf{y} = f(\mathbf{x})$. Letting $\tilde{f}$ denote the inverse function of f, the FO cumulative distribution function (CDF) of RV **y** is given by

$$c_{\mathbf{y}}(y) = \Pr(\mathbf{y} < y) = \Pr(f(\mathbf{x}) < y) = \Pr(\mathbf{x} < \tilde{f}(y)) = c_{\mathbf{x}}(\tilde{f}(y)). \tag{38}$$

The corresponding FO PDF is given by

$$p_{\mathbf{y}}(y) = \frac{d}{dy} c_{\mathbf{y}}(y) = \frac{d}{dy} c_{\mathbf{x}}(\tilde{f}(y)) = p_{\mathbf{x}}(\tilde{f}(y))\, \tilde{f}'(y). \tag{39}$$

But, since

$$y = f(x) = f(\tilde{f}(y)),\ 1 = f'(\tilde{f}(y))\, \tilde{f}'(y), \tag{40}$$

the end result for the FO PDF of RV **y** follows as

$$p_{\mathbf{y}}(y) = \frac{p_{\mathbf{x}}(\tilde{f}(y))}{f'(\tilde{f}(y))}. \tag{41}$$

This last step avoids the need to evaluate the derivative of inverse function $\tilde{f}$, as in equation (39), which is often a much more complicated nonlinear function than f.

As an example, consider the cubic nonlinear transformation

$$y = f(x) = x + a x^3,\ a > 0;\quad f'(x) = 1 + 3 a x^2. \tag{42}$$

Then, the inverse of this transformation is

$$x = \tilde{f}(y) = t(y) - \frac{1}{3 a t(y)}, \tag{43}$$

where

$$t(y) = \left(\frac{y}{2a} + \sqrt{\left(\frac{y}{2a}\right)^2 + \frac{1}{27 a^3}} \right)^{1/3} \quad \text{for all } y. \tag{44}$$

The positive square root is taken to ensure increasing monotonicity with y. Once PDF $p_{\mathbf{x}}(x)$ is specified, all the quantities needed in equation (41) are available.

The detailed solution of the cubic nonlinear transformation in equation (42) proceeds as follows: let $x = t + u$, where t and u are arbitrary. Substitution in equation (42) yields

$$(t+u)(1+3atu) + a(t^3 + u^3) = y. \tag{45}$$

Now let $u = -1/(3at)$, in which case equation (45) reduces to

$$\begin{aligned} & at^3 - \frac{1}{27a^2t^3} = y, \quad t^6 - \frac{y}{a}t^3 - \frac{1}{27a^3} = 0, \\ & t^3 = \frac{y}{2a} + \sqrt{\left(\frac{y}{2a}\right)^2 + \frac{1}{27a^3}}, \quad t = \left(\frac{y}{2a} + \sqrt{\left(\frac{y}{2a}\right)^2 + \frac{1}{27a^3}}\right)^{1/3}, \\ & u = \frac{-1}{3at}, \; x = t - \frac{1}{3at}. \end{aligned} \tag{46}$$

This is the result quoted in equations (43) and (44).

The optimum nonlinearities corresponding to PDF $p_y(y)$ in equation (41) are

$$q(y) = \log p_y(y), \; q_1(y) = q'(y), \; q_2(y) = q''(y). \tag{47}$$

Analytic expressions for these quantities are complicated. Appendix E provides a MATLAB program using symbolic math for these tasks. Simpler analytic fits to $q_1(y)$ and $q_2(y)$ are also included for purposes of significantly reducing the numerical evaluation of equation (47). Figure 7 depicts the optimum nonlinearities and their fits for an example with $a = 0.1$ and a Gaussian RV **x** with zero mean and unity standard deviation.

A comparison was made of the ROCs for several processors, using $K = 1000$ independent data samples and various combinations of nonlinearities q_1 and q_2 in figures 8 through 11. For these simulation results, 1 million trials were used. Processor A used only the linear term $q_1 = \mathbf{r}$ with $q_2 = 0$, where **r** is the received data; processor B used q_1 and q_2 equal to the optimum fits above; processor C used q_1 as the optimum fit and $q_2 = 0$; and processor D used $q_1 = 0$ and q_2 equal to the optimum fit. To achieve false alarm probability P_f = 1e-3 and detection probability $P_d = 0.5$, the required input SNRs for processors A through D were -10.1, -13.3, -13.1, and -12.4 dB, respectively. Instead, to achieve operating point (1e-4, 0.9), the required input SNRs for the four processors were -7.8, -10.8, -10.5, and -9.9 dB, respectively. Finally, for processors B, C, and D, the required input SNRs at operating point (1e-5, 0.99) were -9.2, -8.9, and -8.4 dB, respectively. In all cases, processor B, which used optimum fits q_1 and q_2, performed the best, although processor C trailed by only 0.3 dB; that is, ignoring the contribution of nonlinearity q_2 was not very detrimental. However, processor D, which uses only q_2, was more costly in terms of loss of SNR. It should also be observed that processor A, which used only a purely linear term, suffered a significant 3-dB loss relative to the best,

processor B. Accordingly, it is very important to include the "saturating effect" of the first type of optimum nonlinearity q_1, which tends to suppress the large outliers of the cubed Gaussian process in equation (42).

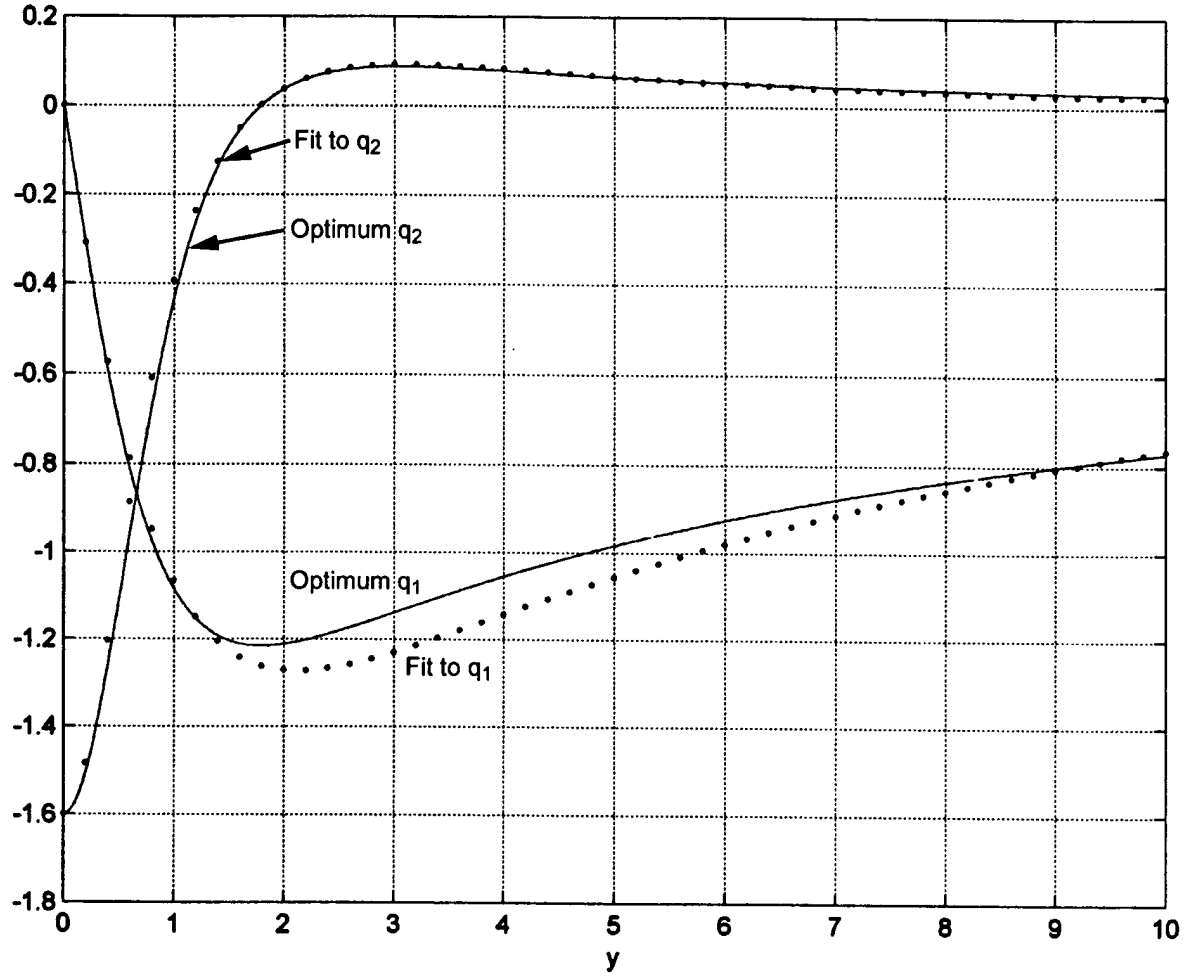


Figure 7. Optimum Nonlinearities and Their Fits

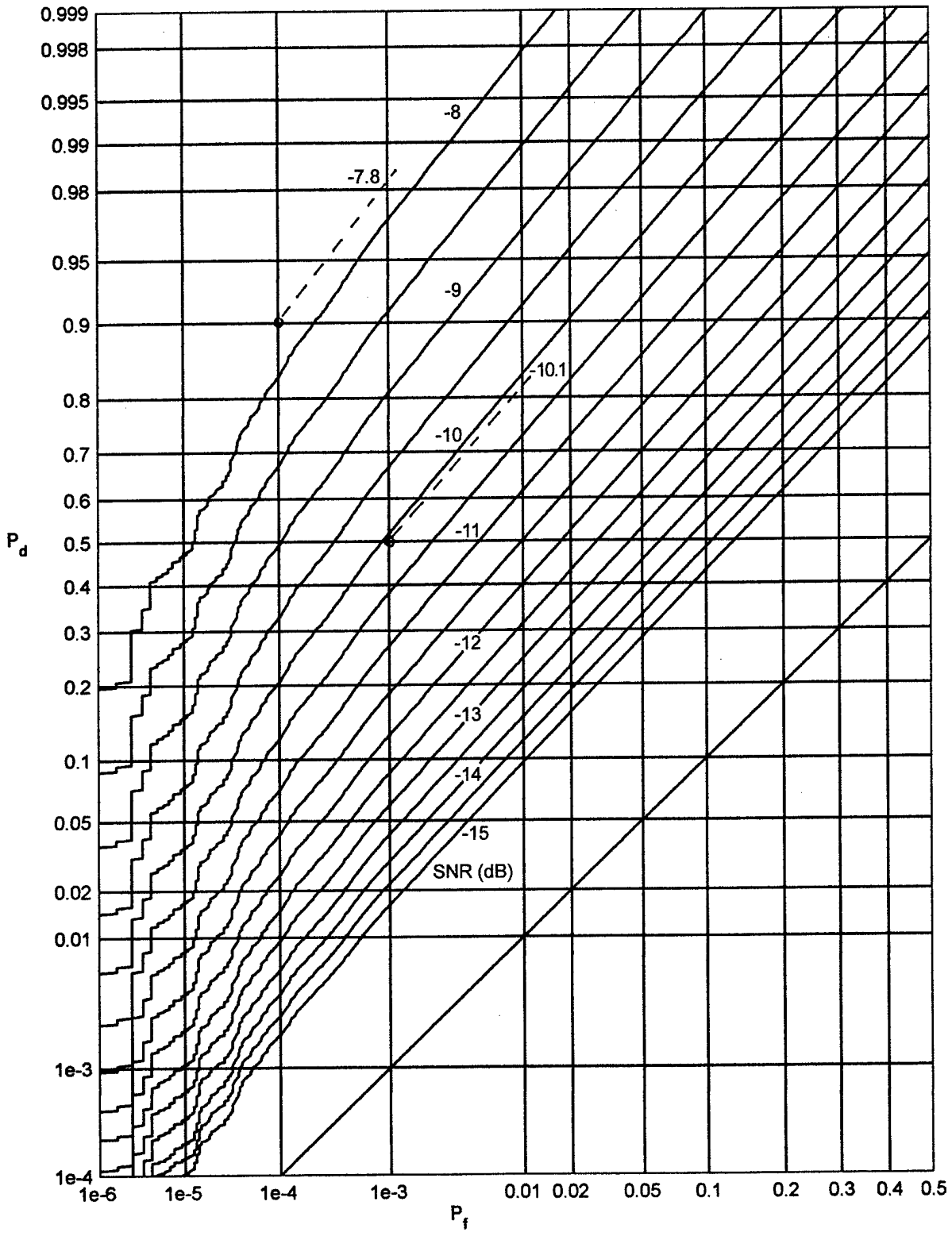


Figure 8. Receiver Operating Characteristics for Processor A

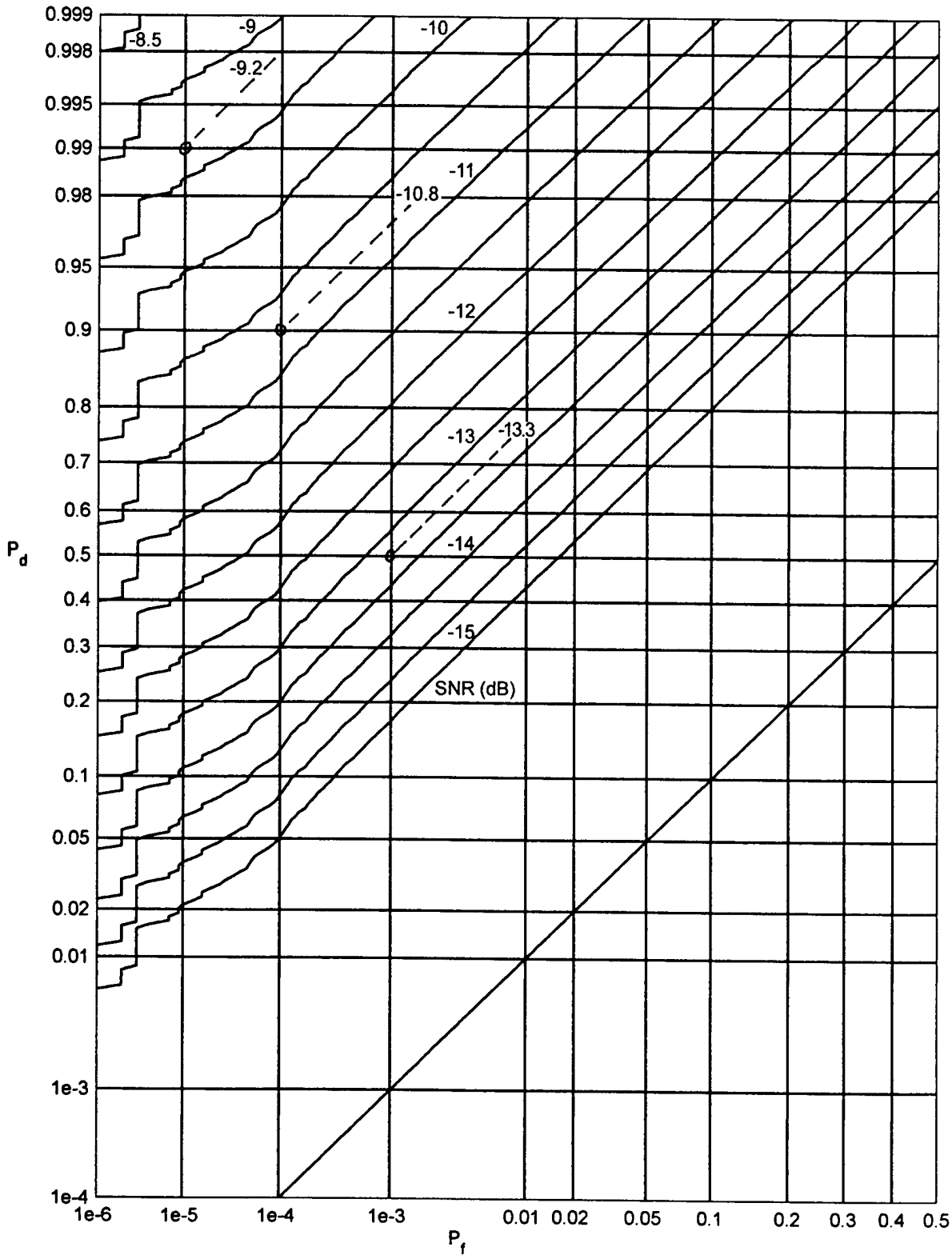


Figure 9. Receiver Operating Characteristics for Processor B

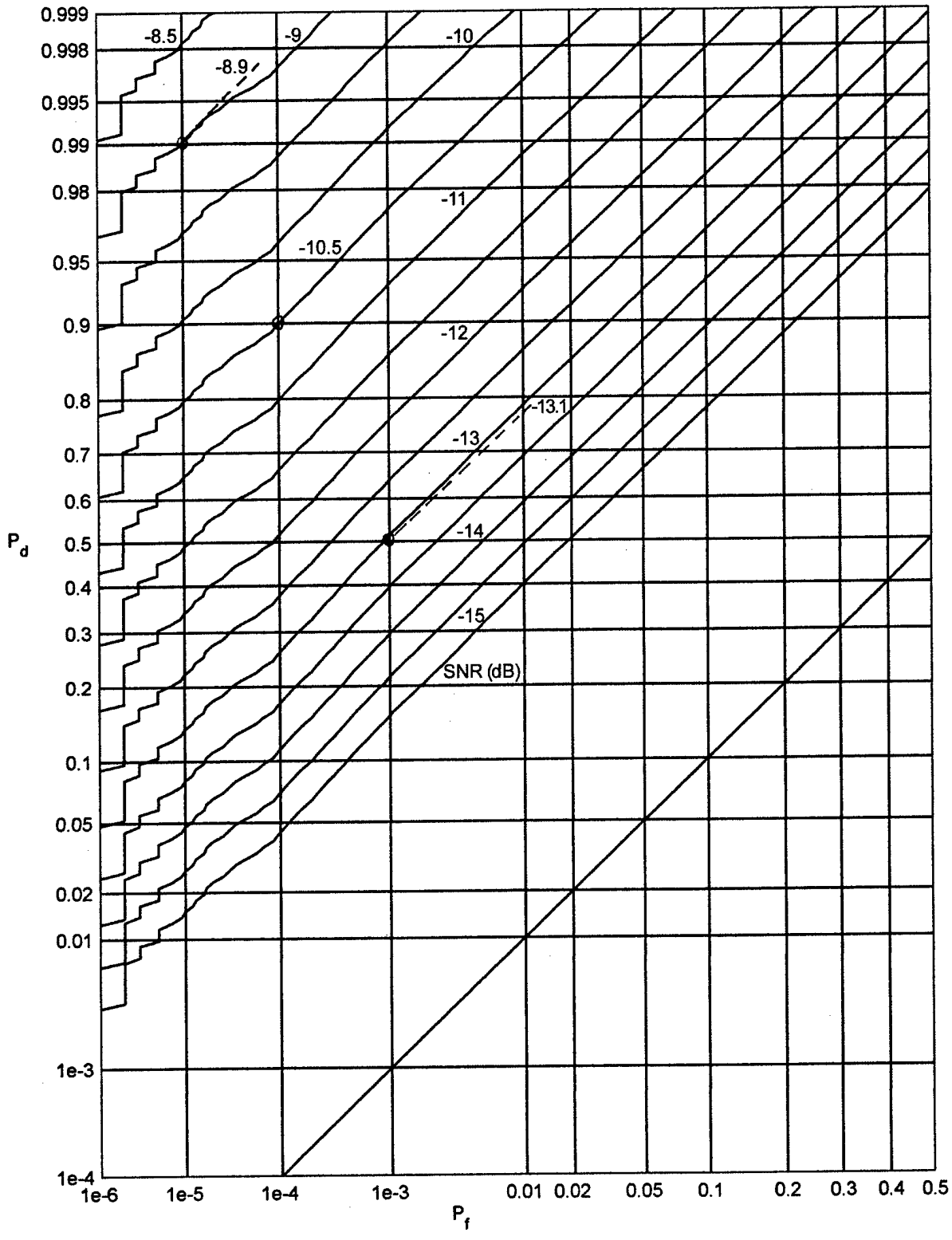


Figure 10. Receiver Operating Characteristics for Processor C

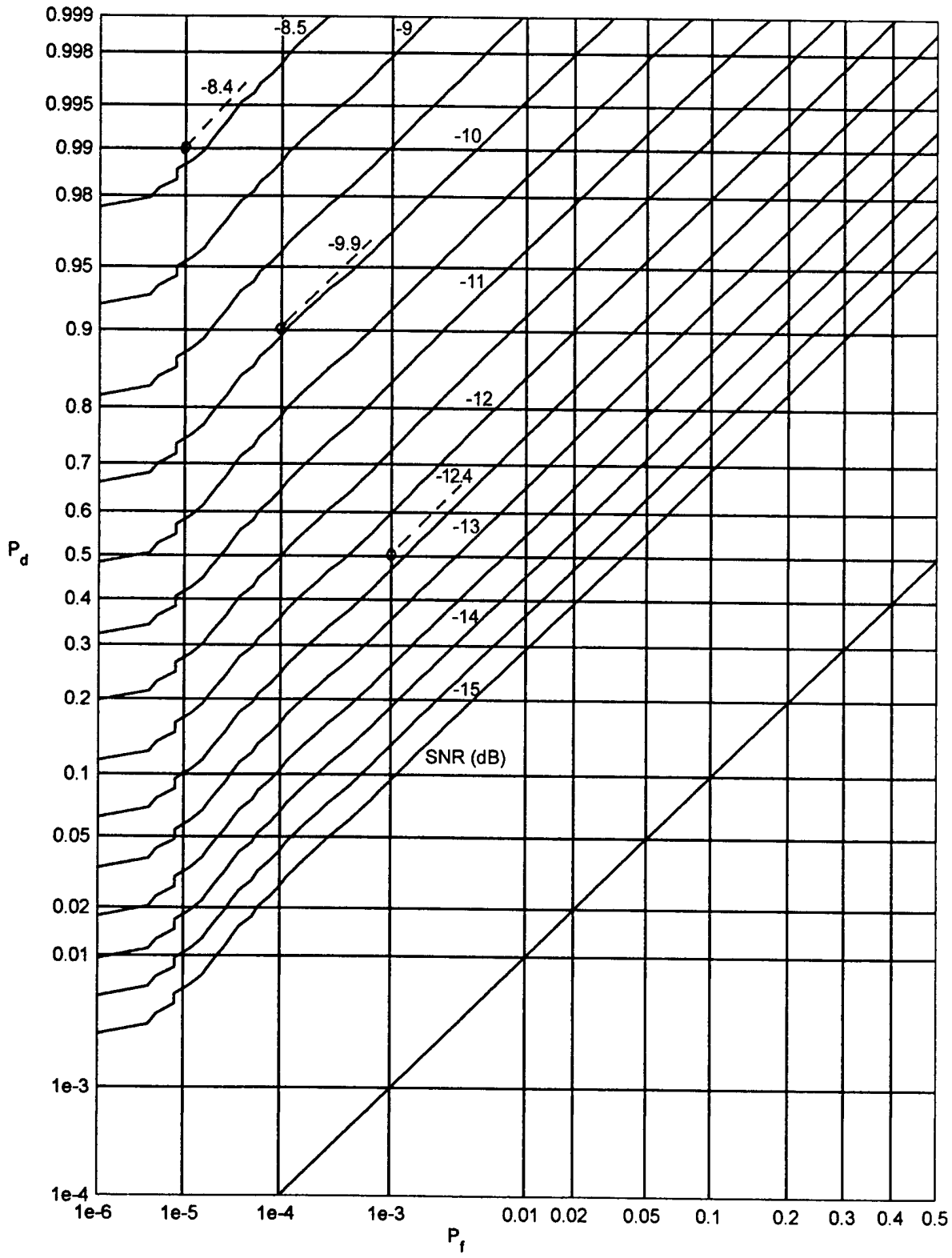


Figure 11. Receiver Operating Characteristics for Processor D

SUMMARY

The optimum processor for a weak stationary random signal in independent non-Gaussian noise has been derived to second order in the signal strength. The information required to realize the likelihood ratio processor consists of the first-order probability density function of the noise process and the covariance function of the signal process. More precisely, the noise density and its first two derivatives are required.

Estimation of these two required derivatives of the noise density, from a finite segment of noise-only data, is not a simple task. The approach adopted here is to fit the noise histogram by means of a spline with several continuous derivatives. Then, analytic derivatives are taken of this fit. For a Gaussian noise process, the second derivative need not be utilized, making this a simpler special case. For a non-Gaussian noise process, both derivatives of the noise density must be estimated, making the estimation task considerably more difficult.

Attempts were made at dropping the second-derivative term in the optimum processor because of its difficulty of reliable estimation. Simulations of the receiver operating characteristics for the corresponding suboptimum processors revealed a loss of several decibels in detectability performance. Accordingly, to achieve optimality it is necessary to conduct the lengthy data collection to enable reliable estimation of the higher-order noise density derivatives and characteristics.

APPENDIX A
EFFECT OF SAMPLING INCREMENT ON PERFORMANCE

The effect of the sampling increment on the deflection criterion in the detection of a deterministic signal in colored noise is of interest. The received waveform, for signal present, is $s(t)+\mathbf{n}(t)$, where $s(t)$ is the known deterministic signal waveform and $\mathbf{n}(t)$ is a stationary noise process. Data processing consists of sampling this received waveform and cross correlating it against the stored reference according to

$$\mathbf{y} = \sum_{k=1}^{K} s(t_k)\,[s(t_k)+\mathbf{n}(t_k)], \tag{A-1}$$

where K is the total number of samples utilized. For zero-mean noise, the mean value of RV $\mathbf{y}$ is, for signal present,

$$\mu(\mathbf{y}) = \sum_{k} s(t_k)^2 . \tag{A-2}$$

The variance of $\mathbf{y}$ is, under both hypotheses,

$$\sigma^2(\mathbf{y}) = \sum_{k,j} s(t_k)\, s(t_j)\, R_n(t_k - t_j), \tag{A-3}$$

where $R_n(\tau)$ is the correlation function of the stationary noise process $\mathbf{n}(t)$. The quantity of interest here is the dimensionless deflection criterion

$$d = \frac{\mu(\mathbf{y})}{\sigma(\mathbf{y})}. \tag{A-4}$$

In the special case of a Gaussian noise process, this statistic completely governs the ROCs; more generally, it is a useful, simple measure of detectability performance.

For equispaced samples with increment Δ, $t_k = k\Delta$, equation (A-4) becomes

$$d = \frac{\sum_{k} s(k\Delta)^2}{\left[\sum_{k,j} s(k\Delta)\, s(j\Delta)\, R_n((k-j)\Delta)\right]^{1/2}}. \tag{A-5}$$

The signal waveform will be taken to be a constant s_o over the *fixed* observation interval of duration T. Then, sampling increment $\Delta = T/K$ and equation (A-5) simplifies to

$$d = \frac{s_o}{\sigma_n} \frac{K}{\left[K + 2\sum_{k=1}^{K-1}(K-k)\,\rho_n(kT/K)\right]^{1/2}}, \tag{A-6}$$

where $R_n(\tau) = \sigma_n^2\ \rho_n(\tau)$, and $\rho_n(\tau)$ is the normalized correlation function of the noise. For example, for white noise over the band (-F,F) of frequencies, equation (A-6) yields

$$d = \frac{s_o}{\sigma_n} \frac{K}{\left[K + 2\sum_{k=1}^{K-1}(K-k)\,\mathrm{sinc}(k\,2FT/K)\right]^{1/2}}, \tag{A-7}$$

where $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$.

As the number of samples K tends to infinity, the second factor in equation (A-7) tends to $(2FT)^{1/2}$ for very large FT. Accordingly, a meaningful quality ratio is furnished by

$$Q \equiv \frac{K/(2FT)^{1/2}}{\left[K + 2\sum_{k=1}^{K-1}(K-k)\,\mathrm{sinc}(k\,2FT/K)\right]^{1/2}}. \tag{A-8}$$

For $K = 2FT$, $Q = 1$. This corresponds to $\Delta = T/K = 1/(2F)$, which is just the sampling increment that results in uncorrelated samples of this white noise process with normalized correlation $\mathrm{sinc}(2F\tau)$. This can be considered to be the standard sampling increment for this particular process. For other values of K, the summation in equation (A-8) must be carried out in numerical detail; the results are presented in figure A-1 for $2FT = 100$. It will be observed that increasing K above $2FT = 100$ does not significantly improve the quality ratio. In fact, K can be decreased to about 60 before a significant loss is observed.

Figure A-2 displays the corresponding quality ratio for a noise process with a Gaussian-shaped spectrum. Again, sampling faster than necessary does not increase the detectability of the signal in noise; the quality ratio simply saturates very quickly once the sampling increment realizes approximately uncorrelated samples. Any higher sampling rate simply requires more data processing with no attendant improvement in detectability performance.

It should be observed that no assumptions have been made about the statistics of the noise process $\mathbf{n}(t)$, such as a Gaussian PDF. The conclusions above apply to stationary noise with arbitrary first-order statistics. It is expected that similar behavior holds for weak random signals (rather than deterministic signals) in noise, when correlators or energy detectors are used for processing.

More generally, as K tends to infinity, the second factor in equation (A-6) tends to

$$\left[\frac{1}{T}\int_{-T}^{T} d\tau \left(1-\frac{|\tau|}{T}\right)\rho_n(\tau)\right]^{-1/2}. \tag{A-9}$$

For very large FT, this quantity approaches

$$\left[\frac{1}{T}\int_{-\infty}^{\infty} d\tau\, \rho_n(\tau)\right]^{-1/2} = \left[\frac{T\,R_n(0)}{G_n(0)}\right]^{1/2}, \tag{A-10}$$

where $G_n(f)$ is the noise spectrum. For the Gaussian-shaped spectrum in figure A-2, the normalized correlation function is $\rho_n(\tau) = \exp(-4\pi F^2\tau^2)$ and the spectrum is $G_n(f) = G_n(0)\exp\left(-\frac{\pi}{4}\frac{f^2}{F^2}\right)$. At frequencies $f = \pm F$, the relative spectral value is down by $\exp(-\pi/4) = 0.456 = -3.41\,\text{dB}$. The second factor in equation (A-6) again tends to $(2FT)^{1/2}$ for large FT, as K increases. The reason that the curves in figures A-1 and A-2 go just above value 1 is due to the factor $(1-|\tau|/T)$ in equation (A-9), which slightly attenuates the integral in the denominator; the limit of 1 is approached only for very large FT.

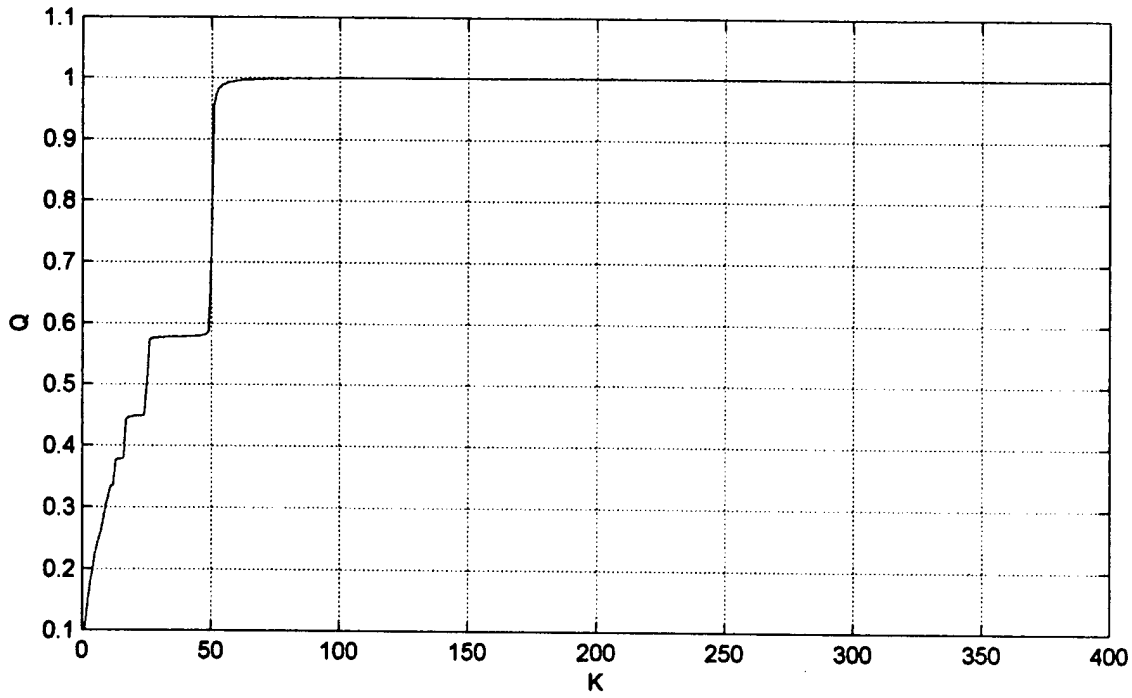


Figure A-1. Quality Ratio for 2FT = 100, Flat Spectrum

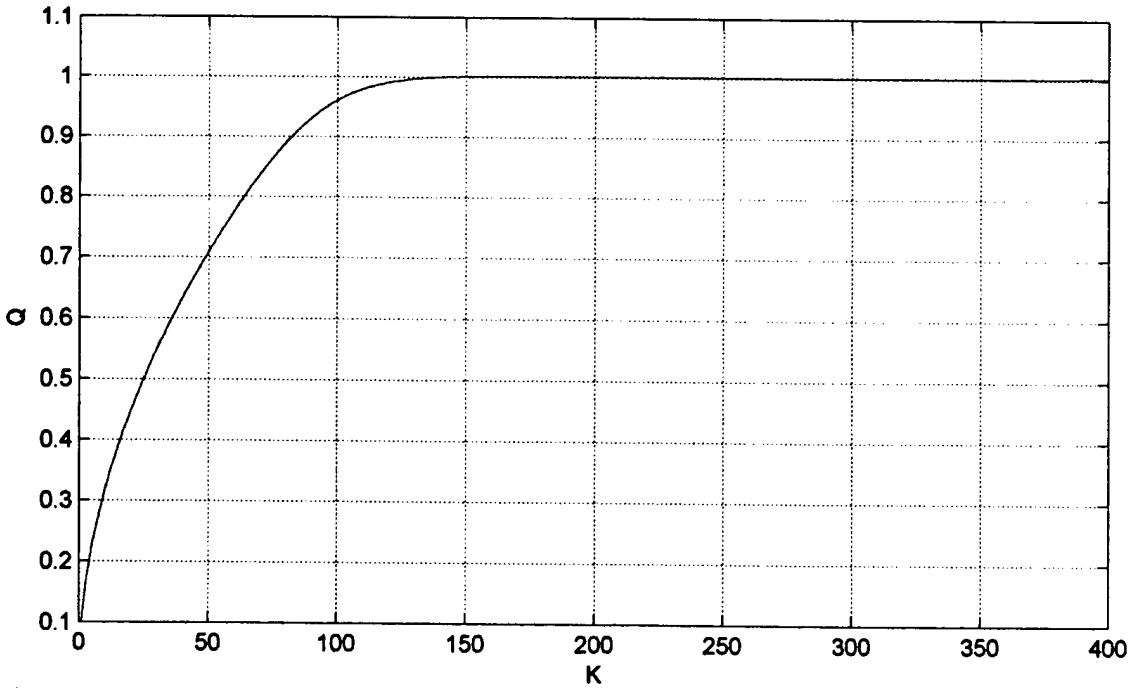


Figure A-2. Quality Ratio for 2FT = 100, Gaussian-Shaped Spectrum

APPENDIX B
HIGHER-ORDER TERMS IN EXPANSION (9)

To determine the third- and fourth-order terms in expansion (9) in the main body of this report, the bracketed expansion (8) is extended to

$$1+u_k\,a(\mathbf{x}_k)+\tfrac{1}{2}u_k^2\,b(\mathbf{x}_k)+\tfrac{1}{6}u_k^3\,c(\mathbf{x}_k)+\tfrac{1}{24}u_k^4\,d(\mathbf{x}_k)\quad \text{for } k=1:K, \tag{B-1}$$

where auxiliary functions

$$c(x)=-\frac{p_n^{(3)}(x)}{p_n(x)},\ \ d(x)=\frac{p_n^{(4)}(x)}{p_n(x)}. \tag{B-2}$$

The product of the K terms of equation (B-1) contains the third-order quantities

$$\tfrac{1}{6}\sum_k c(\mathbf{x}_k)\,u_k^3,\ \ \tfrac{1}{2}\sum_{k\neq j} b(\mathbf{x}_k)\,a(\mathbf{x}_j)\,u_k^2\,u_j,\ \ \sum_{l<k<j} a(\mathbf{x}_l)\,a(\mathbf{x}_k)\,a(\mathbf{x}_j)\,u_l\,u_k\,u_j, \tag{B-3}$$

and the fourth-order quantities

$$\begin{gathered}\tfrac{1}{24}\sum_k d(\mathbf{x}_k)\,u_k^4,\ \ \tfrac{1}{6}\sum_{k\neq j} c(\mathbf{x}_k)\,a(\mathbf{x}_j)\,u_k^3\,u_j,\ \ \tfrac{1}{4}\sum_{k<j} b(\mathbf{x}_k)\,b(\mathbf{x}_j)\,u_k^2\,u_j^2,\\ \tfrac{1}{2}\sum_l b(\mathbf{x}_l)\,u_l^2\sum_{k<j}^{k\neq l} a(\mathbf{x}_k)\,a(\mathbf{x}_j)\,u_k\,u_j,\ \ \sum_{m<l<k<j} a(\mathbf{x}_m)\,a(\mathbf{x}_l)\,a(\mathbf{x}_k)\,a(\mathbf{x}_j)\,u_m\,u_l\,u_k\,u_j.\end{gathered} \tag{B-4}$$

When the quantities in equation (B-3) are substituted in equation (9) in the main report, the averages over the joint signal PDF $p_s(u_1,\ldots,u_K)$ will require knowledge of the general third-order signal moment $E(\mathbf{s}_l\,\mathbf{s}_k\,\mathbf{s}_j)$ for all l,k,j. If the signal samples in equation (3) in the main body of this report are taken from a zero-mean Gaussian process, these third-order moments will all be zero.

On the other hand, the averages over the quantities in equation (B-4) will require knowledge of the general fourth-order signal moment $E(\mathbf{s}_m\,\mathbf{s}_l\,\mathbf{s}_k\,\mathbf{s}_j)$ for all m,l,k,j. For a zero-mean Gaussian signal process, this fourth-order moment can be expressed as combinations of second-order signal covariances coefficients, namely,

$$\sigma_s^4\left(\rho_{ml}\,\rho_{kj}+\rho_{mk}\,\rho_{lj}+\rho_{mj}\,\rho_{lk}\right). \tag{B-5}$$

For other processes, the fourth-order moment is much more difficult to evaluate.

APPENDIX C
MODIFICATION OF APPROXIMATE LIKELIHOOD RATIO PROCESSOR FOR EXPONENTIAL NOISE DENSITY

When the additive noise in model (3) in the main body of this report has an exponential PDF, equation (28) indicates that the $q''(x)$ component of the optimum processor (16) becomes a delta function. This unacceptable term is due to expansion (6) breaking down in the neighborhood of $\mathbf{x}_k = 0$, where the exponential PDF does not have a derivative. A closer look at equation (5) reveals that, for low input SNR, the desired operation is a local smoothing of the joint noise PDF in the neighborhood of the point $\mathbf{x}_1,\ldots,\mathbf{x}_K$. This suggests that nonlinearities $q'(x)$ and $q''(x)$ in equation (28) be smoothed in x by a function of width approximately that of the standard deviation σ_s of the signal. Since the smoothing operation in equation (5) would tend to retain the shape of the noise PDF, the following candidate replacement functions were employed, respectively, instead of $q'(x)$ and $q''(x)$:

$$q_1(x) = -[1-\exp(-|x|/g)]\ \mathrm{sgn}(x), \quad q_2(x) = -\frac{1}{g}\exp(-|x|/g)\,. \tag{C-1}$$

These functions are, respectively, odd and even, with $q_2(x)$ being the derivative of $q_1(x)$, consistent with that same property of original nonlinearities $q'(x)$ and $q''(x)$. The positive scalar g is taken as $c\,\sigma_s$, where c is of the order of 1. Numerical simulations of optimum processor (16), with replacements (C-1), yielded a significant improvement in the ROCs compared with those for the $q'(x)$ nonlinearity in equation (28); the $q''(x)$ nonlinearity in equation (28) had to be completely discarded because it cannot be implemented. A gain of about 4 dB was obtained by use of replacements (C-1) in optimum processor (16) using only $q'(x)$. An additional experiment, in which the $q_2(x)$ term was discarded, led to a degraded performance in the ROCs; accordingly, it is concluded that the $q_2(x)$ term is a significant contributor to the approximate optimum processor and must be retained. The ROC performance using replacements (C-1) was not very sensitive to the exact value employed for scalar c; thus, exact knowledge of the input signal standard deviation σ_s is not a necessity. Also, alternative nonlinearities to those in equation (C-1) were tried, without significant changes or degradation, provided the proper scaling g was employed.

This replacement procedure is an ad hoc approach based upon smoothing the original direct results in equation (28). To lend credence to this approach, an *exact* derivation of the LR processor (4) for a Gaussian signal and exponential noise was conducted for $K = 2$. The exact statistics employed are

$$p_n(x) = \frac{1}{2\beta}\exp\left(-\frac{|x|}{\beta}\right),\ \sigma_n^2 = 2\beta^2,$$
$$p_s(u_1,u_2) = \frac{1}{2\pi\sigma_s^2\left(1-\rho^2\right)^{1/2}}\exp\left(-\frac{u_1^2+u_2^2-2\rho u_1 u_2}{2\sigma_s^2(1-\rho^2)}\right). \tag{C-2}$$

The exact LR is given by equations (4) and (5) with $K = 2$. To evaluate this double integral, the following two integral results are very useful:

$$L(h,k,\rho) = \int_h^\infty dx \int_k^\infty dy\, \frac{1}{2\pi\left(1-\rho^2\right)^{1/2}}\exp\left(-\frac{x^2+y^2-2\rho xy}{2\left(1-\rho^2\right)}\right), \tag{C-3}$$

$$\int_a^\infty dx \int_b^\infty dy\, \exp\left(-\tfrac{1}{2}\alpha x^2 - \tfrac{1}{2}\beta y^2 + \gamma xy + \mu x + \nu y\right)$$
$$= \frac{2\pi}{\sqrt{\mathrm{d}}}\exp\left(\frac{\alpha\nu^2+\beta\mu^2+2\gamma\mu\nu}{2\mathrm{d}}\right) L(\mathrm{u},\mathrm{v},\mathrm{p}), \tag{C-4}$$

where

$$\mathrm{d} = \alpha\beta-\gamma^2,\ u = \sqrt{\frac{\mathrm{d}}{\beta}}\left(a-\frac{\beta\mu+\gamma\nu}{\mathrm{d}}\right),\ \mathrm{v} = \sqrt{\frac{\mathrm{d}}{\alpha}}\left(b-\frac{\alpha\nu+\gamma\mu}{\mathrm{d}}\right),\ \mathrm{p} = \frac{\gamma}{\sqrt{\alpha\beta}}.$$

Definition (C-3) is given in reference 1, section 26.3. The result in equation (C-4) is obtained by making linear shifts of the variables x and y to eliminate the linear variables in the integrand and then scaling to manipulate the double integral into the form of equation (C-3).

Upon substituting equation (C-2) into equation (5) from the main body of this report, splitting the double integral into four parts corresponding to the noise joint-PDF breakpoints at its origin, and using equation (C-4) four times, the following result for the exact LR is obtained for $\sigma_s > 0$ (after considerable manipulations):

$$r = \frac{\sigma_s}{\beta},\ \mathbf{y}_1 = \frac{\mathbf{x}_1}{\beta},\ \mathbf{y}_2 = \frac{\mathbf{x}_2}{\beta},\ \mathbf{a} = |\mathbf{y}_1| + |\mathbf{y}_2|,\ t_p = r(1+\rho),\ t_m = r(1-\rho),$$
$$\mathbf{e}_1 = \exp(r\,t_p + \mathbf{a} - \mathbf{y}_1 - \mathbf{y}_2),\ \mathbf{e}_2 = \exp(r\,t_m + \mathbf{a} + \mathbf{y}_1 - \mathbf{y}_2),$$
$$\mathbf{e}_3 = \exp(r\,t_p + \mathbf{a} + \mathbf{y}_1 + \mathbf{y}_2),\ \mathbf{e}_4 = \exp(r\,t_m + \mathbf{a} - \mathbf{y}_1 + \mathbf{y}_2), \tag{C-5}$$
$$\mathbf{L}_1 = L(t_p - \mathbf{y}_1/r, t_p - \mathbf{y}_2/r, \rho),\ \mathbf{L}_2 = L(t_m + \mathbf{y}_1/r, t_m - \mathbf{y}_2/r, -\rho),$$
$$\mathbf{L}_3 = L(t_p + \mathbf{y}_1/r, t_p + \mathbf{y}_2/r, \rho),\ \mathbf{L}_4 = L(t_m - \mathbf{y}_1/r, t_m + \mathbf{y}_2/r, -\rho),$$
$$LR(\mathbf{x}_1,\mathbf{x}_2) = \mathbf{e}_1\mathbf{L}_1 + \mathbf{e}_2\mathbf{L}_2 + \mathbf{e}_3\mathbf{L}_3 + \mathbf{e}_4\mathbf{L}_4.$$

A computer routine for the joint Gaussian exceedance probability distribution $L(h,k,\rho)$ enables rapid numerical evaluation of this LR for any $\mathbf{x}_1, \mathbf{x}_2$.

The approximate LR processor, obtained from equation (16) by use of replacements (C-1), is given by

$$LR_q(\mathbf{x}_1,\ldots,\mathbf{x}_K) \equiv \sum_{k=1}^{K} q_2(\mathbf{x}_k) + \sum_{k,j=1}^{K} \rho_{kj}\, q_1(\mathbf{x}_k) q_1(\mathbf{x}_j). \tag{C-6}$$

Several comparisons between exact result $LR(\mathbf{x}_1,\mathbf{x}_2)$ and $LR_q(\mathbf{x}_1,\mathbf{x}_2)$ are made in figures C-1 through C-4 for various values of parameters $r, \rho_{12}, c, \mathbf{x}_1/\sigma_n, \mathbf{x}_2/\sigma_n$. Since additive constants and scaling factors do not affect the ROCs of a processor, $LR_q(\mathbf{x}_1,\mathbf{x}_2)$ was transformed so that it had the same values as $LR(\mathbf{x}_1,\mathbf{x}_2)$ at the origin and $+\infty$. The close similarity of the two sets of numerical results indicates that the modified procedure, obtained by smoothing the nonlinearities $q'(x)$ and $q''(x)$ when necessary, is an appropriate and viable procedure; this is especially pertinent when K is large and the K-fold integration of equation (5) is intractable. Then, modification (C-6) becomes an attractive and useful alternative.

It was noted in equation (17) in the main body of this report that the optimum processor conducts a single sum of the transformed input data $\{\mathbf{x}_k\}$ when the input signal samples $\{\mathbf{s}_k\}$ are uncorrelated. However, for exponential input noise, equation (28) reveals that the pertinent nonlinear transformation involves a negative delta function, which is meaningless. Instead, an exact analysis for a Gaussian signal PDF yields LR

$$L(\mathbf{x}_1,\ldots,\mathbf{x}_K) = \prod_{k=1}^{K} \left\{ \begin{array}{l} \exp\left(\tfrac{1}{2}r^2 + |\mathbf{y}_k| + \mathbf{y}_k\right) \Phi(-r - \mathbf{y}_k/r) \\ + \exp\left(\tfrac{1}{2}r^2 + |\mathbf{y}_k| - \mathbf{y}_k\right) \Phi(-r + \mathbf{y}_k/r) \end{array} \right\}, \tag{C-7}$$

where $\Phi(u) = \int_{-\infty}^{u} dt\, \exp(-t^2/2)/\sqrt{2\pi}$ is the normalized Gaussian cumulative distribution function. Plots of the k-th term on the right-hand side of equation (C-7) versus $\mathbf{x}_k$ reveal a function that is a continuous approximation to the negative delta function at the origin, with a width proportional to signal standard deviation σ_s. This is further confirmation of the smoothing procedure that is recommended earlier in this appendix for nonlinearities that are meaningless or cannot be realized in practice. This procedure can be useful for other noise PDFs $p_n(x)$ with non-analytic behavior at some argument(s).

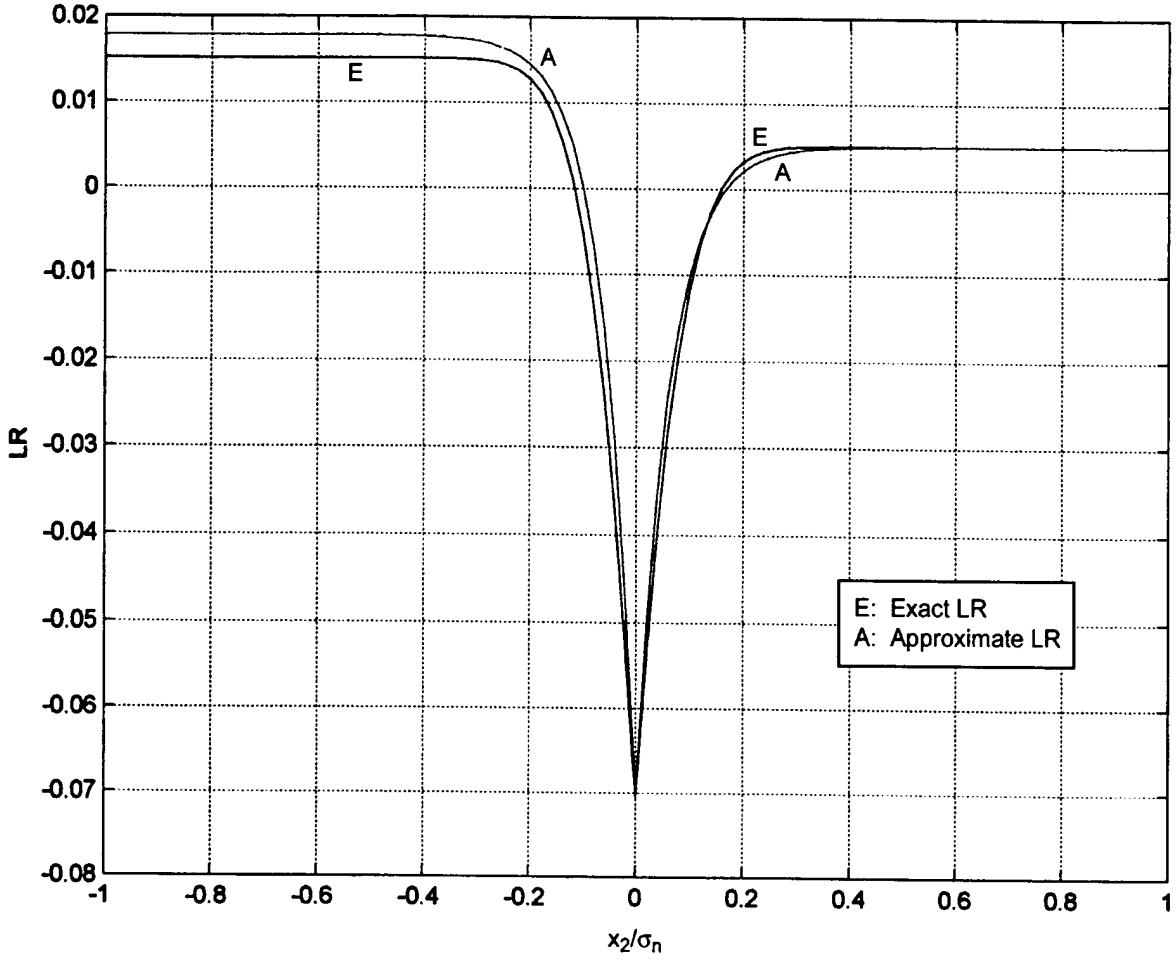


Figure C-1. Likelihood Ratios for $r = 0.1$, $\rho_{12} = 0.5$, $c = 0.6$, $x_1/\sigma_n = -0.6$

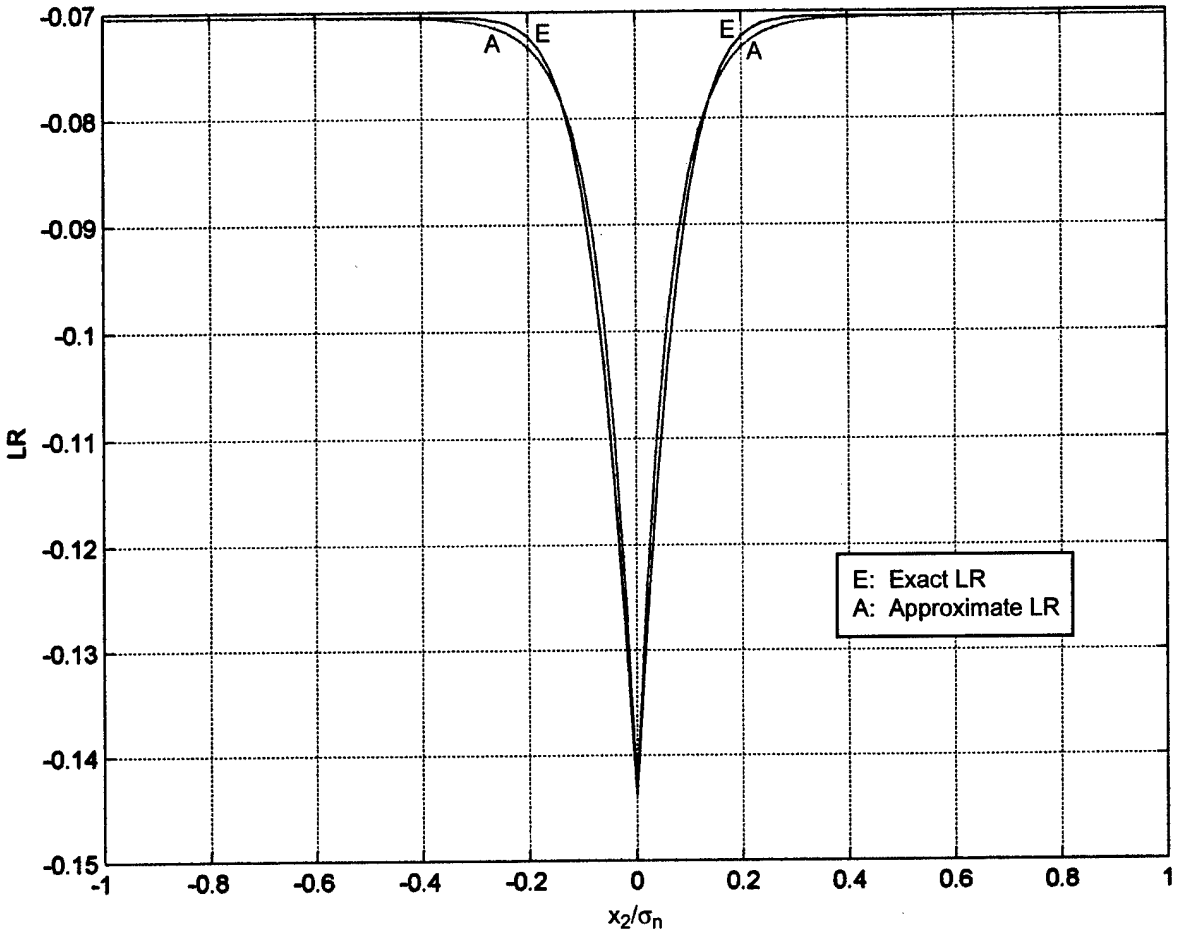


Figure C-2. Likelihood Ratios for $r = 0.1$, $\rho_{12} = 0.5$, $c = 0.6$, $x_1/\sigma_n = 0$

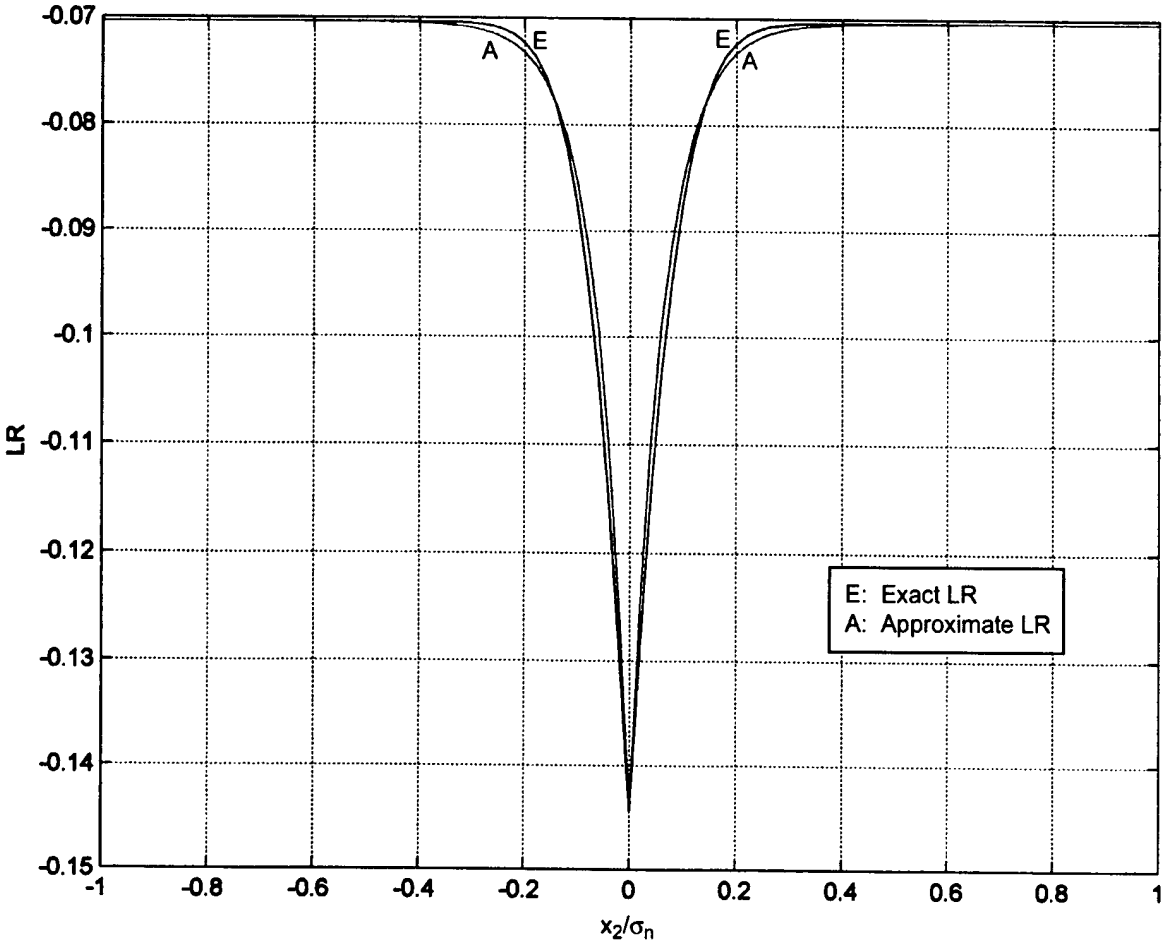


Figure C-3. Likelihood Ratios for $r = 0.1$, $\rho_{12} = 0.0$, $c = 0.6$, $x_1/\sigma_n = 0$

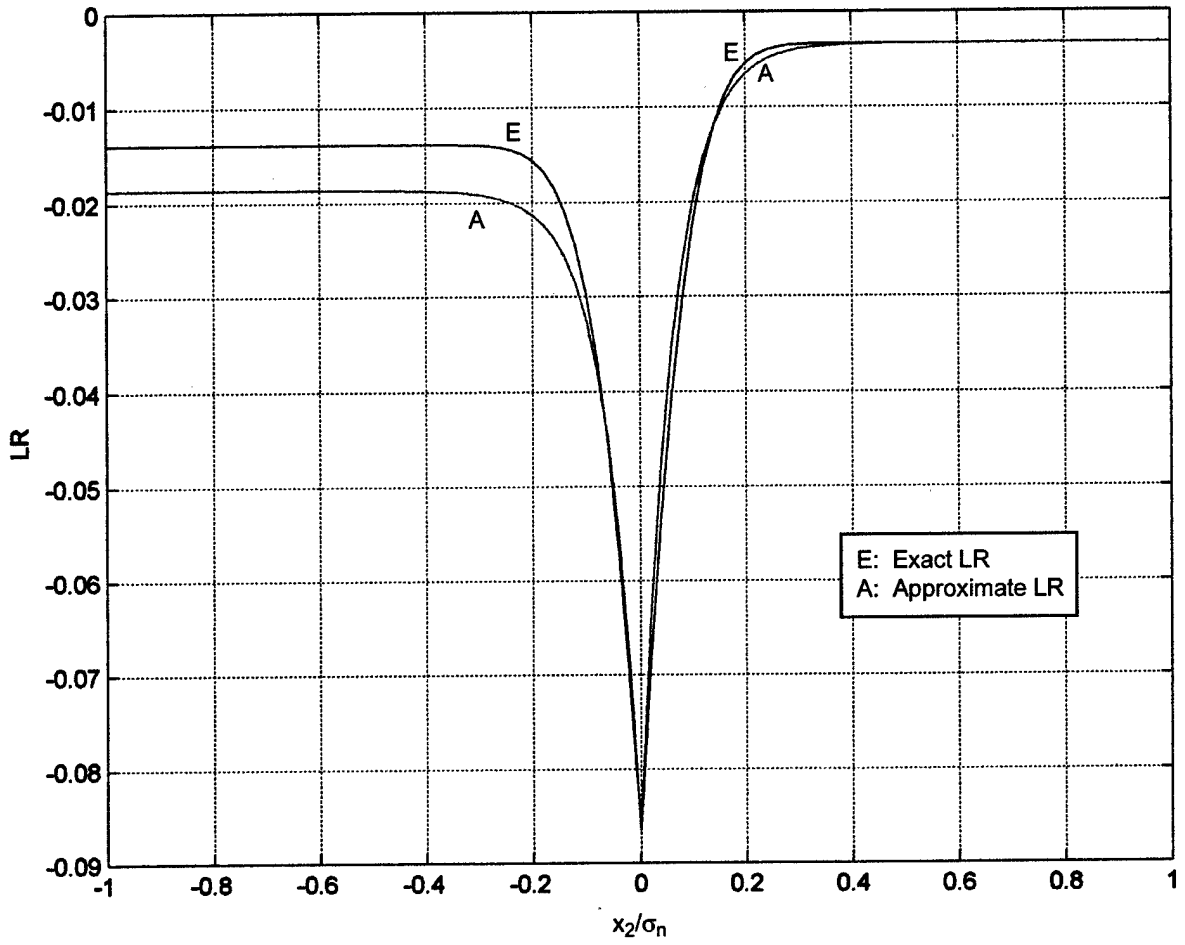


Figure C-4. Likelihood Ratios for $r = 0.1, \rho_{12} = 0.8, c = 0.6, x_1/\sigma_n = 0.1$

APPENDIX D
TWO IMPULSIVE NOISE MODELS

The first noise model is the square of an exponential RV, followed by a multiplication by an equally likely ±1. This RV can be generated by the operation

$$\mathbf{u} = \pm\left(\frac{1}{b}\log(\mathbf{r})\right)^2, \quad b = \frac{(24)^{1/4}}{\sigma^{1/2}}, \tag{D-1}$$

where RV **r** is uniformly distributed over (0,1), and σ is the desired standard deviation of RV **u**. The PDF of RV **u** is

$$p(u) = \frac{b}{4}\frac{\exp(-b\sqrt{|u|})}{\sqrt{|u|}} \quad \text{for all } u, \tag{D-2}$$

and its even moments are given by

$$E\{\mathbf{u}^{2n}\} = \frac{(4n)!}{24^n}\sigma^{2n}. \tag{D-3}$$

The odd moments of **u** are zero.

The EDF of **u** is given by

$$e(u) = \begin{Bmatrix} \frac{1}{2}\exp(-b\sqrt{u}) & \text{for } u > 0 \\ 1-\frac{1}{2}\exp(-b\sqrt{|u|}) & \text{for } u < 0 \end{Bmatrix}. \tag{D-4}$$

The characteristic function of **u** is

$$f(\xi) = \int du\, \exp(i\xi u)\, p(u) = \sqrt{\pi}\ \mathrm{Re}\left\{\frac{(1-i)b}{\sqrt{8|\xi|}}\, w\left(\frac{(1+i)b}{\sqrt{8|\xi|}}\right)\right\} \quad \text{for } \xi \neq 0. \tag{D-5}$$

This is a closed form in terms of the $w(z)$ function (see reference 1, chapter 7).

The second noise model is the fourth power of a zero-mean Gaussian RV, followed by a multiplication by an equally likely ±1. Letting **q** be a ±1 RV with equal probabilities, for an arbitrary RV **x**, the RV $\mathbf{y} = \mathbf{q}\,\mathbf{x}$ has the following properties:

$$\begin{aligned} f_y(\xi \mid \mathbf{q}) &= E_{\mathbf{x}}\{\exp(i\xi\mathbf{q}\mathbf{x})\} = f_x(\xi\mathbf{q}), \\ f_y(\xi) &= \tfrac{1}{2} f_x(\xi) + \tfrac{1}{2} f_x(-\xi) = \mathrm{Re}\, f_x(\xi), \\ p_y(u) &= \tfrac{1}{2} p_x(u) + \tfrac{1}{2} p_x(-u) \quad \text{for all } u, \\ e_y(u) &= \tfrac{1}{2} e_x(u) + \tfrac{1}{2} c_x(-u) \quad \text{for all } u. \end{aligned} \tag{D-6}$$

For the particular case of RV $\mathbf{x} = \mathbf{g}^4$, where **g** is a zero-mean unit-variance Gaussian RV, the following functions govern the statistics of RVs **x** and **y**:

$$\begin{aligned} f_x(\xi) &= \int du\, \frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2}u^2 + i\xi u^4\right) = \frac{1+i}{\sqrt{\pi 32\xi}} \exp\left(\frac{i}{32\xi}\right) K_{1/4}\left(\frac{i}{32\xi}\right) \quad \text{for } \xi > 0, \\ f_y(\xi) &= \mathrm{Re}\left\{\frac{1+i}{\sqrt{\pi 32|\xi|}} \exp\left(\frac{i}{32|\xi|}\right) K_{1/4}\left(\frac{i}{32|\xi|}\right)\right\} \quad \text{for } \xi \neq 0, \quad f_y(0) = 1, \\ p_x(u) &= \frac{1}{4u^{3/4}}\, \phi(u^{1/4}) \quad \text{for } u > 0; \quad e_x(u) = 2\,\Phi(-u^{1/4}) \quad \text{for } u > 0, \\ p_y(u) &= \frac{1}{8|u|^{3/4}}\, \phi(|u|^{1/4}) = \frac{1}{\sqrt{2\pi}\, 8|u|^{3/4}} \exp(-\tfrac{1}{2}\sqrt{|u|}) \quad \text{for all } u \neq 0, \\ e_y(u) &= \begin{Bmatrix} \Phi(-u^{1/4}) & \text{for } u > 0 \\ \Phi(|u|^{1/4}) & \text{for } u < 0 \end{Bmatrix}. \end{aligned} \tag{D-7}$$

The variance of RV **y** is 105. All the odd moments are zero. The functions

$$\phi(x) = (2\pi)^{-1/2} \exp(-x^2/2), \quad \Phi(x) = \int_{-\infty}^{x} dt\, \phi(t), \tag{D-8}$$

are, respectively, the PDF and cumulative distribution function of a normalized Gaussian RV.

APPENDIX E
PROGRAM FOR OPTIMUM q FUNCTIONS

```
clear all                                   % Optimum q1(y) and q2(y)
syms y a pi                                 %  for Y = X + a X^3
sx=1;                                       % Standard deviation of X

t1=y/(2*a);
b=1/(3*a);
t=(t1+sqrt(t1^2+b^3))^(1/3);
x=simplify(t-b/t);                          % Solution of y = f(x) = x + a x^3
fp=1+3*a*x^2;                               % f'(x)
px=exp(-.5*(x/sx)^2)/sqrt(2*pi*sx^2);       % Gauss PDF(x) of X
py=simplify(px/fp);                         % PDF(y) of Y = f(X)
q=log(py);
q1=diff(q,y);                               % Optimum q1(y)
q1=simplify(q1);
q1i=inline(q1);                             % q1i(a,y)
q2=diff(q1,y);                              % Optimum q2(y)
q2=simplify(q2);
q2i=inline(q2);                             % q2i(a,y)

ad=.1;                                      % a = 0.1 example
yd=[0:.02:10]';
q1d=q1i(ad,yd);
figure(1), clf
plot(yd,q1d,'r')
grid on, hold on
q2d=q2i(ad,yd);
plot(yd,q2d,'b')

yd=[0:.2:10]';
f1d=-1.6*yd./(1+.5*yd.^1.6);                % Fit to q1(y) for ad=.1
plot(yd,f1d,'r.')                           % Fit to q2(y) for ad=.1
f2d=1.6*((yd/1.79).^2-1)./(1+1.6*yd.^2+.2*yd.^4);
plot(yd,f2d,'b.')
xlabel('y')
```

REFERENCES

1. *Handbook of Mathematical Functions*, U.S. Department of Commerce, National Bureau of Standards, Applied Mathematics Series, No. 55, U.S. Government Printing Office, Washington, DC, June 1964.

INITIAL DISTRIBUTION LIST

Addressee	No. of Copies
Office of Naval Research (ONR 32 (D. Abraham, B. Fitch, D. Johnson, F. Herr, J. Tague))	5
Naval Sea Systems Command (PEO-IWS 5 (R. Zarnich))	1
University of Connecticut (P. Willett)	1
University of Rhode Island (S. Kay)	1
Defense Technical Information Center	2
Center for Naval Analyses	1